



**Earth Sciences**  
New Zealand

# **Water quality state and trends in New Zealand lakes**

Analysis of national lakes data ending in 2024

*Prepared for Ministry for the Environment*

**October 2025**

Prepared by:

Doug Booker, Rachel Smith, David Wood, Caroline Fraser (LWP), Ton Snelder (LWP)

For any information regarding this report please contact:



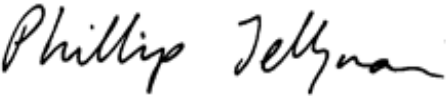
David Wood  
Water Quality Scientist  
Freshwater Modelling Group  
+64 3 343 8050  
david.wood@niwa.co.nz

New Zealand Institute for Earth Science Limited  
PO Box 8602  
Riccarton  
Christchurch 8440

Phone +64 3 348 8987

Client Report No: 2025334CH  
Report date: October 2025  
Project No: MFE25504

Revision	Description	Date
Version 1.0	Final version sent to client	24 October 2025

Quality Assurance Statement		
	Reviewed by:	Bruce Dudley
	Formatting checked by:	Nic McNeil
	Approved for release by:	Phillip Jellyman

---

© New Zealand Institute of Earth Science Limited (“Earth Sciences New Zealand”) 2025. All rights reserved. This publication may not be reproduced or copied in any form without the permission of the copyright owner(s). Such permission is only to be given in accordance with the terms of the client’s contract with Earth Sciences New Zealand. This copyright extends to all forms of copying and any storage of material in any kind of information retrieval system.

Whilst Earth Sciences New Zealand has used all reasonable endeavours to ensure that the information contained in this document is accurate, Earth Sciences New Zealand does not give any express or implied warranty as to the completeness of the information contained herein, or that it will be suitable for any purpose(s) other than those specifically contemplated during the project or agreed by Earth Sciences New Zealand and the client.

# Contents

- Executive summary ..... 7**
  - Purpose .....7
  - Methods .....7
  - Results .....8
  - Discussion.....9
  
- 1 Introduction .....10**
  - 1.1 Background .....10
  - 1.2 Study purpose.....10
  - 1.3 General approach .....10
  
- 2 Data .....12**
  - 2.1 Water quality variables .....12
  - 2.2 Censoring information.....13
  - 2.3 Water quality data .....13
  - 2.4 Quality control information.....14
  - 2.5 Measurement method information .....15
  - 2.6 Landscape-scale data .....16
  - 2.7 Water quality data processing .....16
  
- 3 Analysis methods.....20**
  - 3.1 Water quality state analyses .....20
  - 3.2 Water quality trend analyses .....25
  
- 4 Results – lake state .....36**
  - 4.1 State summary statistics .....36
  - 4.2 NOF grades.....45
  - 4.3 Relationships between water quality state and catchment land cover .....52
  
- 5 Results – lake trends .....54**
  - 5.1 Ten-year trends (2015–2024).....54
  - 5.2 Twenty-year trends (2005–2024) .....62
  - 5.3 Thirty-year trends (1995–2024).....68
  - 5.4 Rolling ten-year trends .....74

<b>6</b>	<b>Discussion .....</b>	<b>79</b>
6.1	Factors to consider when interpreting results .....	79
6.2	Additional analyses and interpretive approaches not applied in this report ...	82
6.3	Differences to previous studies.....	83
6.4	Overview.....	84
<b>7</b>	<b>Acknowledgements.....</b>	<b>85</b>
<b>8</b>	<b>References .....</b>	<b>86</b>
<b>Appendix A</b>	<b>Output data .....</b>	<b>89</b>

## Tables

Table 2-1:	Lake water quality variables included in this study.	12
Table 2-2:	NEMS QC codes and their named quality zones (source: NEMS 2019).	14
Table 3-1:	Details of the NOF attributes used to grade the state of lake monitoring sites.	22
Table 3-2:	Level of confidence categories used to convey the confidence that the trend direction was improving.	33
Table 4-1:	Number of lakes (and sites, in brackets) by geomorphology class and water quality variable included in the state analyses.	36
Table 4-2:	Number of lakes (and sites, in brackets) by percentage of catchment human modified land cover and water quality variable included in the state analyses.	37
Table 4-3:	Summary of the number of lakes and sites (in brackets, where this is different from the number of lakes) in each NOF grade.	46
Table 5-1:	Number of lakes and sites (in brackets where different from the number of lakes) in the lake geomorphological classes included in the 10-year trend analyses.	54
Table 5-2:	Number of lakes and sites (in brackets where different from the number of lakes) in human modified land cover classes included in the 10-year trend analyses.	55
Table 5-3:	Proportions of improving 10-year lake trends (P <sub>i</sub> ) by variable.	61
Table 5-4:	Number of lakes and sites (in brackets where different from the number of lakes) in the lake geomorphology classes included in the 20-year trend analyses.	62
Table 5-5:	Number of lakes and sites (in brackets where different from the number of lakes) in human modified land cover classes included in the 20-year trend analyses.	62
Table 5-6:	Proportions of improving trends (P <sub>i</sub> ) for 20-year time period.	67
Table 5-7:	Number of lakes and sites (in brackets where different from the number of lakes) in the lake geomorphology classes included in the 30-year trend analyses.	68

Table 5-8:	Number of lakes and sites (in brackets where different from the number of lakes) in human modified land cover classes included in the 30-year trend analyses.	68
Table 5-9:	Proportions of improving trends (P <sub>i</sub> ) for 30-year time period.	73

## Figures

Figure 3-1:	Schematic diagram demonstrating how the Kendall S statistic and confidence in trend direction (C) is calculated.	29
Figure 3-2:	Schematic diagram of the calculation of the Sen slope, which is used to characterise trend rate.	31
Figure 4-1:	Locations of monitored lakes and number of sites per lake used for state analyses of lake water quality.	38
Figure 4-2:	Proportion of censored data for each site × variable combination for which state was assessed.	39
Figure 4-3:	For each variable, the proportion of analysed data that was not labelled QC 200.	39
Figure 4-4:	Lake water quality state in lake geomorphology classes.	41
Figure 4-5:	Lake water quality state in human modified land cover classes.	42
Figure 4-6:	<i>E. coli</i> concentrations in lake geomorphological classes.	43
Figure 4-7:	<i>E. coli</i> concentrations in human modified land cover classes.	44
Figure 4-8:	<i>E. coli</i> percent exceedance in lake geomorphological classes.	44
Figure 4-9:	<i>E. coli</i> percent exceedance in human modified land cover classes.	45
Figure 4-10:	Stacked bar chart showing the percentage of lakes assigned each NOF grade overall NOF attribute.	46
Figure 4-11:	Maps showing lake site NOF overall attribute grades.	47
Figure 4-12:	Maps showing lake site grades for NOF attribute components.	48
Figure 4-13:	Stacked bar charts showing the percentage of lakes assigned to each NOF grade), by lake geomorphological class for NOF overall attributes.	49
Figure 4-14:	Stacked bar charts showing the percentage of lakes assigned to each NOF grade (median over sites within a lake), by lake upstream human modified land cover classes for NOF overall attributes.	50
Figure 4-15:	Stacked bar charts showing the percentage of lakes assigned to each NOF grade (median over sites within a lake), by lake geomorphological class for NOF attribute components.	51
Figure 4-16:	Stacked bar charts showing the percentage of lakes assigned to each NOF grade (median over sites within a lake), by lake upstream human modified land cover classes for NOF attribute components.	52
Figure 4-17:	Relationships between lake median water quality state and proportion of upstream catchment with human modified land cover.	53
Figure 5-1:	Locations of lake sites included in the 10-year trend analyses by water quality variable.	56
Figure 5-2:	Proportion of censored data for each site × variable combination for which trend was assessed.	57
Figure 5-3:	Summary of 10-year lake trend rates for the water quality variables, in lake geomorphological classes.	58

Figure 5-4:	Summary of 10-year lake trend rates for the water quality variables in human modified land cover classes.	59
Figure 5-5:	Lake water quality monitoring sites categorised by the confidence that the 10-year trend was improving ( $C_i$ ) for each variable.	60
Figure 5-6:	Summary plot representing the proportion of lakes with improving 10-year trends at each categorical level of confidence.	61
Figure 5-7:	Locations of lake sites included in the 20-year trend analyses by water quality variable.	63
Figure 5-8:	Summary of 20-year lake trend rates for the water quality variables, in lake geomorphological classes.	64
Figure 5-9:	Summary of 20-year lake trend rates for the water quality variables in human modified land cover classes.	65
Figure 5-10:	Lake water quality monitoring sites categorised by the confidence that the 20-year trend was improving ( $C_i$ ) for each variable.	66
Figure 5-11:	Summary plot representing the proportion of lakes with improving 20-year time-period trends at each categorical level of confidence.	67
Figure 5-12:	Locations of lake sites included in the 30-year trend analyses by water quality variable.	69
Figure 5-13:	Summary of 30-year lake trend rates in water quality variables, in lake geomorphological classes.	70
Figure 5-14:	Summary of 30-year lake trend rates in water quality variables in human modified land cover classes.	71
Figure 5-15:	Lake water quality monitoring sites categorised by the confidence that the 30-year trend was decreasing ( $C_i$ ) for each variable.	72
Figure 5-16:	Summary plot representing the proportion of lakes with improving 30-year time-period trends at each categorical level of confidence.	73
Figure 5-17:	Number of lakes included for each 10-year period in the rolling trends assessment by variable.	74
Figure 5-18:	Proportion of censored data for each site $\times$ variable combination for which 10-year rolling trends were assessed.	75
Figure 5-19:	Summary of rolling window 10-year trend rates (relative Sen slopes, %).	76
Figure 5-20:	Summary plot representing the proportion of lakes with improving 10-year time-period trends at each categorical level of confidence for each time window.	77
Figure 5-21:	Summary plot representing the proportion of improving site trends ( $P_i$ ) for each of the rolling 10-year windows.	78

## Executive summary

### Purpose

The New Zealand Ministry for the Environment (MfE) and Stats NZ Tatauranga Aotearoa (Stats NZ) use results from national-scale analyses of lake water quality originating from state-of-the-environment (SoE) monitoring to inform policy development and meet their requirements for environmental reporting on the freshwater domain under the Environmental Reporting Act 2015. The current study analysed lake water quality variables representing physical, chemical, microbial, and ecological conditions observed at lake surfaces across multiple monitoring sites for the period ending in December 2024. Following conventions used in previous national environmental reporting and in consultation with MfE, nine monitoring variables were analysed for current state and temporal trend: Secchi depth (Secchi), concentrations of nitrate-nitrite-nitrogen (NNN), ammoniacal nitrogen (NH<sub>4</sub>-N), total nitrogen (TN), dissolved reactive phosphorus (DRP), total phosphorus (TP), phytoplankton biomass as chlorophyll *a*, the faecal bacterium *Escherichia coli* (*E. coli*), and Trophic Level Index (TLI). Various organisations monitor these nine water quality variables across many lakes for various purposes. A large volume of observation and associated metadata therefore exist, but are sometimes held in different formats, different units, and using different data conventions. The purpose of this report was collation, standardisation, and preliminary analysis of available water quality data for subsequent use in environmental reporting.

### Methods

A data request was made directly to all regional councils and unitary authorities to obtain all water quality data collected routinely for SoE purposes from commencement of monitoring to 31 December 2024. Relevant censoring information, observation depths, site metadata, and descriptions of measurement methods were also requested. Questions that arose after initial data collation about data quality, completeness, or purpose were resolved using subsequent queries to data providers.

Data were collated into a common format to ensure consistent conventions for variable names, site identifiers, date and time formats, units of measurement, the depth relating to observations, and other data structure elements. Transcription errors, incomplete censoring information, or potentially erroneous values were identified and resolved following data inspection and additional queries with data providers. Predefined rules for data exclusion were applied based on information supplied by data providers such as data quality, non-SoE monitoring purposes, or non-comparable data collection methods where they were known. Site metadata were used to collate information describing the upstream catchment characteristics for each lake. The upstream proportion of aggregated agricultural, urban, and exotic forest land cover was used to represent a gradient of human modified land cover.

Duration and frequency of observations for each site × variable combination were assessed against predetermined criteria for inclusion in statistical analysis. For several variables, some observations were provided as censored values because they were reported as a value less than an analytical detection limit or greater than a reporting limit. Predefined methods were used to consistently replace censored values by imputation prior to analysis of state and robust methods were used to handle censored values during analysis of trends.

We assessed state by calculating summary statistics for each site × variable combination within a 5-year period ending December 2024. Site medians were grouped by four evenly split classes of human modified land cover. We also grouped sites by geomorphology class. Linear regressions were used to inspect the relationship between median water quality state and human modified land cover. Site × variable combinations were assigned to a predefined attribute state defined by bands labelled A to E.

The trend rate for 10-, 20- and 30-year periods ending in December 2024 was assessed using the Sen slope estimator. The seasonal version of the Sen slope estimator was applied for variables measured seasonally (i.e., monthly, bi-monthly or quarterly), and where within-site seasonal variability was detected. Kendall tests provided estimates of the confidence in the trend direction. Trends for all site × variable combinations were classified using predefined categories describing the confidence that a given trend was improving. Temporal transience of trends was evaluated for each variable at each site using rolling windows of 10-years duration starting in 1990 and incrementing by one year to a final window ending December 2024.

Two approaches were used to evaluate aggregate trends in each water quality variable at the national scale and within human modified land cover classes. The first approach tallied the proportion of lakes in each category of trend direction and confidence. The second approach used the confidence in trend direction and confidence for individual sites to estimate the proportion of all sites with improving trends and its 95% confidence interval.

## Results

Sufficient data for calculation of state was available for between 79 and 123 lakes (including 102 to 154 sites), depending on water quality variable. Lakes were distributed across a gradient of human modified land cover and across the country. Many lakes were not included in our analysis because they are not being monitored for water quality, or the available data did not meet criteria for inclusion in our analysis.

Water quality cannot be measured precisely when observed data are below an analytical detection limit (left-censored) or above a reporting limit (right-censored). For chlorophyll *a*, Secchi, TN, and TP, many sites had a small proportion of censored observations during the state assessment period, and sites that had a large proportion of censored observations were rare. DRP, NH<sub>4</sub>-N, *E. coli*, and NNN exhibited a larger proportion of censored observations at more sites compared to the other variables. TLI data are not censored.

Attribute states for chlorophyll *a*, TN, and TP, at many sites (60, 48, 37% respectively) were graded D or E. For the *E. coli* overall numeric attribute state, 8% of sites were graded D or E.

Variation in state for nutrient and *E. coli* concentrations and Secchi was partly distinguished by human modified land cover classes. However, state of *E. coli* concentrations and Secchi varied considerably within lakes at either end of the human modified land cover gradient. Median concentrations of each nutrient (except for NNN) increased, and Secchi decreased, with increasing proportions of human modified land cover in the upstream catchment.

For 10-year trends, all variables were improving for some lakes and worsening at other lakes. It was difficult to discern general patterns in lake trends due to the relatively small number of lakes for which trend could be calculated. However, for chlorophyll *a*, TLI, DRP, and NH<sub>4</sub>-N, a majority of lakes were worsening. Improving and worsening of water quality were not systematically associated with the proportion of human modified land cover in the upstream

catchment. A different set of lakes were used to calculate 10-, 20-, and 30-year trends due to more sites being available in recent years. Results for 20-year trends indicated that for chlorophyll *a* and *E. coli*, a majority of lakes were worsening, and for DRP, TP, and NH<sub>4</sub>-N a majority of lakes were improving.

Analysis of rolling 10-year trend windows provided insights into changes in aggregate trends over time. The national-scale proportions of improving trends were not constant or monotonic for any variable. Quasi-periodic fluctuations in proportions of improving trends were particularly strong for Secchi, DRP, and TP, and less strong for chlorophyll *a* and TN.

## Discussion

The statistical power of state and trend analyses and the degree to which lakes in the analyses represented all lakes in New Zealand were limited by the small number of lakes with sufficient data for analysis. The small numbers of lakes resulted from the scarcity of lakes in council SoE monitoring networks and the exclusion of some monitored lakes due to inadequate data.

We produced state and trend results by applying a prescribed set of methodological procedures to available water quality SoE data. We then discuss the possibility that results are sensitive to methodological choices (e.g., analysis period), monitoring network design (e.g., site positions and longevity), and observation data features (e.g., detection limits, data precision, observation frequency). One example is the confidence in trend direction at a single site being sensitive to the proportion of censored data and changes in data precision. A second example is the proportion of improving trends calculated for different time periods being sensitive to duration of analysis period and closing or opening of sites between periods. This report explains how our analysis applied robust procedures to mitigate, or explore, some undesirable but unavoidable characteristics of SoE data. For example, transient decreases and increases within rolling trends of water quality data indicated that caution must be applied when interpreting causes and permanence of calculated trends.

We explain that the results presented in this study differ from those produced for previous studies of water quality state and trend because the present study used up-to-date data, improved methods of data preparation (e.g., filtering by quality control codes), use of NNN instead of nitrate-nitrogen (NO<sub>3</sub>-N), different methods for grouping lakes, and improved statistical methods.

# 1 Introduction

## 1.1 Background

State-of-environment (SoE) monitoring for lake water quality involves collecting, analysing, and reporting on physical, chemical, and ecological data to assess health and trends over time. SoE monitoring involves measuring a suite of standard variables at representative sites to provide a comprehensive and unbiased understanding of water quality condition (state) and its changes (trends).

Lake water quality data measured for SoE purposes is typically collected from strategically selected sites within lakes at regular, but relatively infrequent, intervals for long durations (e.g., monthly for a duration of years). Lake water quality data can also be measured at SoE sites or elsewhere at frequent intervals for short durations (e.g., daily or hourly for a single month) for non-SoE purposes such as bespoke scientific studies, validation or calibration of instrumentation, or to issue public health warnings.

Monthly or quarterly monitoring has been carried out at approximately 115 SoE lake sites across New Zealand by New Zealand local government authorities including regional councils and unitary councils, hereafter referred to as “councils”. This monitoring has been carried out for at least five years and continues to the present. A variety of physical, chemical and biological indicators of water quality (referred to in this report as “variables”) are measured at these sites.

The New Zealand Ministry for the Environment (MfE) and Stats NZ Tatauranga Aotearoa (Stats NZ) use analyses of state and trends derived from SoE lake water quality data to inform policy development, and to meet their requirements for environmental reporting on the freshwater domain under the Environmental Reporting Act 2015. In this report, we use “lake water quality” as a general term to refer to physical, chemical, and biological variables that are often measured as part of lake SoE monitoring programmes.

## 1.2 Study purpose

Stats NZ are required to produce up-to-date environmental statistics in the form of environmental reporting indicators under the Environmental Reporting Act 2015 (ERA). Stats NZ relies on access to data collected by various organisations to subsequently produce surface water quality indicators required by the ERA. Under the ERA, Stats NZ decides what statistics will be used to measure environmental reporting topics. Under the Data and Statistics Act 2022, Stats NZ also has the sole responsibility for deciding on the production of statistics by the Statistician, including the selection of methods. Under both Acts, Stats NZ also has a duty to act independently. In addition to the requirements of Stats NZ, MfE use results from national-scale analyses of river water quality originating from SoE monitoring to inform freshwater policy development. The purpose of this report was collation, standardisation, and preliminary analysis of available lake water quality data for subsequent use in environmental reporting and to inform freshwater policy.

## 1.3 General approach

Council lake monitoring data have been periodically acquired and compiled into databases for preparation of national-scale SoE reports and to investigate monitoring performance. Earlier reports for MfE provided lake water quality state and trends based on monitoring data with ending dates of December 2017 (Larned et al. 2018) and December 2020 (Whitehead et al.

2021). For the current report, we newly acquired all lake water quality data. To assess data completeness, we compared the coverage of the newly acquired data to that used for the preceding national-scale report (Whitehead et al. 2021). We then undertook systematic analyses to report updated states and trends to cover the period up to December 2024. We communicated results using graphs, tables, and statistical summaries. We provided all raw results as appendices to this report. The detailed methods and tabulated, site-specific results will enable MfE to use the results for a wide range of purposes (e.g., mapping, inter-comparisons between environmental classes or geographic domains, estimation of reference conditions) that are all based on a single comprehensive methodology.

The main components of the current report are detailed methods for data processing and analysis, summaries of water quality state and trends at the national scale and within four classes representing a gradient in human modified land cover, and supplementary files with site-specific results and spatial data for each site.

This work consisted of several major steps:

1. Compile SoE lake water quality monitoring data along with associated quality QC information, site metadata, censoring information, observation depths, and measurement methods from regional councils.
2. Post-process and collate the data into a national database that uses a common format, including error correction, application of reporting conventions, classification of surface and sub-surface observations, and links to lake spatial data.
3. Assess whether there is sufficient data for statistical analyses for each of nine physical, chemical, microbial, and ecological variables by applying site inclusion rules.
4. Carry out analyses of water quality state for many lakes, including comparison with numeric attribute states defined in the National Policy Statement for Freshwater Management 2020.
5. Investigate patterns in the water quality state with regard to catchment characteristics by aggregation of lake water quality state by human modified land cover classes, and graphing relationships with a gradient of human modified land cover.
6. Carry out trend analyses for 10-, 20- and 30-year periods ending in December 2024 for as many lakes as possible, including comparisons of trends aggregated by land-cover classes.
7. Evaluate trends for each variable at each site for rolling windows of 10-years duration starting in 1990 and incrementing by one year (ending 31 December) to a final window ending in 2024 (i.e., a total time period of 35 years).
8. Assess water quality trends at the national scale using categorical levels of confidence and a statistical analysis of the proportions of improving trends.

## 2 Data

In this section we describe the water quality variables, data sources, and organisation of the lake water quality database, and explain the data processing procedures used to derive datasets suitable for state and trend analyses.

### 2.1 Water quality variables

We assessed lake water quality using nine variables that correspond to physical, chemical and biological conditions (Table 2-1). The lake water quality variables were Secchi depth, concentrations of ammoniacal nitrogen, oxidised (or nitrate-) nitrogen, total nitrogen, dissolved reactive phosphorus, total phosphorus, chlorophyll *a* (a photosynthetic pigment used as an index of phytoplankton biomass), and the Trophic Level Index. Hereafter, the variables are referred to by the abbreviations listed in Table 2-1.

**Table 2-1: Lake water quality variables included in this study.**

Variable type	Variable	Abbreviation	Units
Physical	Secchi depth	Secchi	m
	Chlorophyll <i>a</i>	Chlorophyll <i>a</i>	mg L <sup>-1</sup>
Biological	Trophic Level Index	TLI	unitless
	<i>Escherichia coli</i>	<i>E. coli</i>	cfu or MPN 100 mL <sup>-1</sup>
Chemical	Dissolved reactive phosphorus	DRP	mg L <sup>-1</sup>
	Total phosphorus (unfiltered)	TP	mg L <sup>-1</sup>
	Ammoniacal nitrogen	NH <sub>4</sub> -N	mg L <sup>-1</sup>
	Nitrate + nitrite nitrogen	NNN	mg L <sup>-1</sup>
	Total nitrogen (unfiltered)	TN	mg L <sup>-1</sup>

Secchi depth (abbreviated as Secchi) is a measure of water clarity and gives an indication of the amount of light-scattering and light-absorbing particulate and dissolved matter in lakes. Secchi measures the maximum depth at which a black and white Secchi disk is visible to an observer at the lake surface.

Five different nutrient species (NNN, NH<sub>4</sub>-N, DRP, TN and TP) were included because they influence the growth of planktonic, epiphytic and benthic algae and vascular plants (macrophytes) in lakes, and because ammonia can be toxic to lake organisms at high concentrations. Nutrient enrichment can promote proliferations of planktonic algae (phytoplankton) and epiphytic algae on the surfaces of lake macrophytes. These algae can inhibit macrophyte growth by reducing light penetration. At elevated concentrations, free ammonia (NH<sub>3</sub>) can be toxic to animals, including lake fish and invertebrates (Randall and Tsui 2002). The concentration of free ammonia and consequent risk to fish and invertebrates is determined by water temperature, pH and salinity, as well as the concentration of total ammonium plus ammonia (NH<sub>4</sub><sup>+</sup> + NH<sub>3</sub>), in this report termed ammoniacal nitrogen and abbreviated NH<sub>4</sub>-N.

Chlorophyll *a* concentration is an index of lake phytoplankton biomass. High chlorophyll *a* concentrations may occur during periods of high internal and/or external nutrient loading, and are the primary indicators of eutrophication. Phytoplankton chlorophyll *a* concentrations are also used to calculate Trophic Level Index scores, as described below.

The Trophic Level Index (TLI) is an indicator variable that summarises data related to lake trophic state and potential primary production. The TLI is used to classify New Zealand lakes into trophic classes (e.g., oligotrophic, eutrophic); TLI scores increase with increasing eutrophication. There are two versions of TLI in use in New Zealand, one with three variables (TLI3) and one with four variables (TLI4) (Burns et al. 2000; Verburg et al. 2010). TLI3 scores are derived from log-transformed concentrations of chlorophyll *a*, TN and TP. TLI4 uses Secchi data in addition to chlorophyll *a*, TN and TP concentrations. However, Secchi data were not available for all lakes in the current study. Moreover, Secchi data are strongly influenced by factors that are independent of trophic state, such as fine glacial sediment and tannins. To ensure consistent calculations, we calculated both TLI3 and TLI4 scores for all lakes in our national dataset (using the formulae given by Burns et al. 1999) and used these scores in lieu of TLI scores provided in council datasets.

The concentration of the bacterium *Escherichia coli* (*E. coli*) is used as an indicator of human or animal faecal contamination, which is associated with the risk to humans arising from infection or illness from waterborne pathogens during contact-recreation.

We used attributes for lakes that have been defined by the NPS-FM to provide context to the water quality state analyses. Five of the nine variables used in the current report are also attributes in the NPS-FM: phytoplankton (as chlorophyll *a*), TN, TP, NH<sub>4</sub>-N and *E. coli*.

## 2.2 Censoring information

For several water quality variables, some true values are too low or too high to be measured with precision. For very low values of a variable, the minimum acceptable precision corresponds to the analytical "detection limit" (DL) for that variable; for very high values of a variable, the minimum acceptable precision corresponds to the "reporting limit" (RL) for that variable. Cases where values of variables are below the detection limit or above the reporting limit are often indicated by the data entries "<DL" (i.e., below detection limit) and ">RL" (i.e., above reporting limit) respectively, where DL and RL are the laboratory detection limit and reporting limit, respectively. Water quality datasets from New Zealand lakes often include DRP, TP, and NH<sub>4</sub>-N measurements that are below detection limits, and *E. coli* and Secchi measurements that are above reporting limits. Replacement of censored values with constant multiples of the detection and reporting limits can result in misleading results when statistical tests are subsequently applied to those data (Helsel 2012). Changes in laboratory procedures or equipment can impact observed water quality time-series through changes in DL, RL, and general data precision. Changes in DL and RL therefore need to be acknowledged within trend assessment methods. We quantified the proportion of censored observations for each site × variable combination for each of our analyses. In this study, repeatable procedures were used to handle censored data in the state (see Section 3.1.2) and trend (see Section 3.2.4) analyses.

## 2.3 Water quality data

Lake water quality monitoring data have been acquired periodically from regional councils for previous national scale analyses for MfE and for scientific studies (e.g., Whitehead et al. 2021;

Kuczynski et al. 2024). For each successive analysis, data were used to update or regenerate a database comprising site information, sampling dates, and measurements for several monitoring variables. Previously derived databases also contained metadata (e.g., methods, alternative variable labels, analytical detection limits, depth of sampling).

We used traceable procedures to acquire lake water quality observation data. MfE first contacted councils to notify them of the data request and to request a suitable contact for each data provider. We then sent the same data request to all data providers. Data and information describing observed water quality, censoring information, site metadata, depth of sampling, and descriptions of measurement methods were requested. We received data directly from data providers via email. We used the data acquired through these procedures to generate a dataset similar in format to that used for the previous national-scale analysis.

After initial data collation, we determined that some data providers had pre-filtered their data to supply data solely collected at regular intervals at SoE monitoring sites within a lake for SoE monitoring purposes, whereas other providers had supplied all available water quality data regardless of purpose or frequency of data collection. We therefore requested that data providers supply information that identified the purpose for which each observation was collected or confirm that data supplied to us were solely associated with routine SoE monitoring rather than for other purposes. We used the information supplied to us by data providers regarding the purpose of data collection to maintain spatial and temporal consistency within our analysis by limiting the analysis to SoE data. For example, we removed frequent observations made for non-SoE purposes to maintain consistency between sites experiencing standard SoE sampling regimes and sites experiencing irregular sampling. We also removed observations made at non-SoE sites to maintain consistency between regions who supplied data only at SoE sites and regions who supplied all data including from bespoke scientific investigations.

## 2.4 Quality control information

We aimed to acquire a dataset covering the period up to 2024 that was broadly comparable to the dataset covering the period up to 2020 that was collated by Whitehead et al. (2021). Quality Control (QC) standards and codes have been applied by data providers since implementation of National Environmental Monitoring Standards (NEMS) QC standards in 2019 (NEMS 2019). We therefore attempted to acquire QC information to allow data labelled as being of poor quality to be identified and removed from our analysis. We requested that data providers supplied NEMS QC codes (Table 2-2) alongside each observed water quality value.

**Table 2-2: NEMS QC codes and their named quality zones (source: NEMS 2019).**

QC code	Named quality zone
QC 600	Highest quality code
QC 500	Fair
QC 400	Compromised
QC 300	Synthetic
QC 200	No quality

QC code	Named quality zone
QC 100	Missing record
QC 0	Non verified

## 2.5 Measurement method information

Lake water quality measurements can be made with a variety of laboratory and field methods. These methods have evolved over time, driven by changes in technology, instrumentation, test availability, economic considerations, and evolving monitoring requirements, such as the need to detect attributes at low levels.

The NEMS aim to ensure that water sample collection, measurement, and data processing are consistent and reliable across New Zealand. This includes standardised monitoring, including the choice of measurement methods, to ensure robust, high-quality and comparable water quality monitoring and recording results between regional councils.

Though different measurement methods are intended to give comparable results, systematic differences in observed values can arise from changes in measurement methods (Davies-Colley and McBride 2016). Different methods being applied to monitor the same variable at different sites could result in systematic between-site or between-region differences in the estimated state. “Step changes” in values associated with changes in method at a site may confound assessment of trend (Wood 2024).

We requested information about laboratory methods as part of our data requests. Our intention was to investigate possible influences of measurement method within our analysis and only retain observations measured using a standard set of methods where necessary. We received some useful information about measurement methods from data providers. One example of an inconsistency in measurement method was that some councils estimated TN using the indirect method<sup>1</sup>, which involves summing up the measurements of TKN and NNN, rather than the direct method<sup>2</sup>, alkaline persulfate digestion. However, descriptions of measurement methods supplied to us by data providers used a wide range of naming conventions and most method data did not give precise descriptions of which methods had been applied at which sites through time. The information we received was therefore not sufficiently comprehensive to determine a measurement method for each observation contained within our database. Given the available information about measurement methods, we were unable to explicitly account for unknown differences between methods within our analysis.

The use of filtered samples for water quality variables labelled TN and TP is a potential issue raised in previous reports (e.g., Whitehead et al. 2021). If samples are filtered, we would expect these variables to be labelled as total dissolved nitrogen (TDN) or total dissolved phosphorus (TDP) or similar instead of TN and TP, and these samples would provide non-comparable results (Horowitz 2013). The method information we received was therefore not sufficiently

<sup>1</sup> Total Nitrogen – Indirect (TN-K) - Calculated from the measurement of Total Kjeldahl Nitrogen (TKN, measured via APHA 4500-N<sub>org</sub>D) + NNN. This method may be preferable when TSS is high.

<sup>2</sup> Total Nitrogen – Direct (TN-A) - APHA 4500-N C OR APHA 4500-P J potassium persulfate digestion then analysis by APHA 4500-NO<sub>3</sub>I. This method is preferable for low-nutrient river waters, having a lower detection limit than the TN-K.

comprehensive to consistently identify the presence of mis-labelled filtered “TN” and “TP” samples across water quality observations and data providers, and we assume variables are correctly labelled and thus that TN and TP are unfiltered. We note that unfiltered samples are used for TN and TP methods described in [NEMS 2019](#).

## 2.6 Landscape-scale data

Booker and Wilkins (2025) describe a complete and comprehensive set of lake data layers for the full and up-to-date representation of all lakes (excluding those classified as Artificial Constructed and Mine) represented by the National Lake Polygons Layer 2024 (NLPL), which is also known as Freshwater Ecosystems of New Zealand (FENZ 2024). For example, they describe calculating land cover conditions in catchments upstream of lakes using Land Cover Database Version 6.0 (LCDB6 as supplied to us by MfE) which comprises proportional cover of 33 land-cover classes, generated from satellite imagery collected in 2023. LCDB6 data were projected onto the lake catchments extracted from a DEM to calculate the proportion of each land cover class occupying the catchment upstream of each lake.

Whitehead et al. (2021) aggregated lakes by lake elevation × depth classes. We applied an alternative method recommended by Fraser and Snelder (2021) that retains the simplicity of land cover categories but avoids the need to assign a dominant land cover, which is to use a categorical subdivision of a gradient in land cover that represents human modified land cover. A gradient in human modified land cover was defined by adding the proportion of catchment area upstream of each lake occupied by Pastoral (LCDB6 classes 40, 30, 33), Urban (LCDB6 classes 1, 2, 5, 6), and Exotic Forest (LCDB6 classes 71, 64) land cover types and subdividing this gradient into four evenly spaced categories.

We applied an additional method of aggregating lakes by geomorphology class from the FENZ lakes database (Leathwick et al. 2010; [weblink](#)) following the approach of (Kuczynski et al. 2024). We did not use the geomorphology classification in the updated FENZ database (FENZ 2024) because the greater number of classes would lead to small numbers of lakes in each class. Further, the feature of the geological classification that allows lakes to belong to multiple geomorphology classes in the geomorphology classification of FENZ 2024 created ambiguity of how to group lakes in these cases.

## 2.7 Water quality data processing

Lake water quality data were processed to ensure that data collated from different sources were represented in a consistent format, that the data were as error-free as possible given available information, that consistent measurement procedures were used where these were known, and that site information was complete and accurate.

Development of the data processing steps required additional queries to data providers to ensure issues were resolved appropriately. Common issues included incomplete censoring information, missing data that were present in previous rounds of reporting, missing site metadata, clarification regarding monitoring purpose, and clarification regarding potentially erroneous values such as zero or negative values.

### Step 1. Reporting conventions.

The water quality data received from data providers varied in reporting formats, reporting conventions for variable names, site identifiers, date and time formats, units of measurement,

and other data structure elements. We first organised data from all sources into a single format. Then we applied a consistent set of reporting conventions for water quality variable names, sample date, observed values, units, censor information, detection limits (where present), sampling depth (where present), quality control code, and site identifier. Where water quality variable names were ambiguous or did not match the variables we requested, we sought clarification with data providers. Observed values were converted to equivalent units where inconsistent measurements units were indicated (e.g., from  $\mu\text{g L}^{-1}$  to  $\text{mg L}^{-1}$ ). In cases where data providers indicated that observed values were raw with no censoring applied, we determined the detection limit associated with each raw value and applied censoring to raw values lower than the detection limit. Quality control codes were represented as NEMS codes where possible. Where no quality control code was present, or a code system could not be equated to a NEMS code, then QC 200 was used. We resolved any inconsistencies in spelling or capitalisation (e.g., transcription errors) of site identifiers to ensure internal consistency. We used time-series plots and quantile plots to identify potential data transcription errors and missing censoring information. Sample depth information, where available, was used to classify observations into “surface” or “sub-surface” observations. Sample depth information included depth classes (e.g., SU-surface, BT-bottom), descriptors (e.g., “75%depth”, “@ bottom”, “top sample taken”), or numeric depth values for which observations at the shallowest sample were used as “surface”. Where no sample depth information was available, we assumed that samples were taken from the surface.

## Step 2. Calculation of composite and adjusted observations.

Where appropriate, composite water quality variables were calculated by summing separate constituent components from the raw observed data. NNN was calculated as the sum of  $\text{NO}_2\text{-N}$  and  $\text{NO}_3\text{-N}$  where they were available and where a NNN observation was not already present. TN was calculated as the sum of TKN and NNN for Waikato Regional Council data, for which TN data was not provided. The effect of censoring of constituent components was considered when they were summed to calculate a composite observation. If only one constituent component was censored because it was at or below the detection limit, then the value of the censored constituent was halved before it was added to the other constituent. If both constituent components were censored, then the raw constituent components were summed and the calculated composite observation was flagged as being at or below its detection limit. The lower quality code of the constituents was associated with the newly calculated composite observation.

We applied a pH correction to  $\text{NH}_4\text{-N}$  to adjust values to equivalent pH 8 values, following the methodology outlined in Hickey (2014). For pH values outside the range of the correction relationship (pH 6–9), the maximum (pH<6) and minimum (pH>9) correction ratios were applied. pH adjustment of ammonia was performed after imputation of censored values (Section 3.1.2). In results tables and figures adjusted ammoniacal-N is abbreviated as “ $\text{NH}_4\text{-N (Adj.)}$ ”.

TLI was calculated for each limnological year (1-Sep to 31-Aug) from monthly, bimonthly or quarterly samples, where samples were present and met minimum data requirements (3 of 4 quarters, 5 of 6 bimonths, or 10 of 12 months) for the required indicators (TN, TP, chlorophyll *a*, Secchi). Calculation of TLI was performed after imputation of censored values (see Section 3.1.2). TLI4 was calculated from annual mean TN, TP, chlorophyll *a*, and Secchi following Burns (1999), and TLI3 was calculated from annual mean TN, TP and chlorophyll *a*.

### Step 3. Data exclusion.

We excluded data which was not relevant or it was not possible to use in our analyses. Data not relevant to our analysis included water quality variables not included in those we requested (e.g., horizontal clarity tube), data for areas not included in our request (i.e., Chatham Islands), data collected for non-SoE purposes, and sub-surface observations. We received data for sites on the Chatham Islands, but these data were excluded from our analysis because they have not been included in previous environmental reporting (Whitehead et al. 2022) and landscape-scale data (i.e., catchment and land cover) is not available for them. Data that could not be used in our analysis included observations with missing observed values, dates, or associated variable name. We excluded overlapping data where we could identify that it was provided by more than one data provider or had been supplied as overlapping datasets from a single data provider.

We further excluded data identified as potentially erroneous. Zero and negative values were excluded, following censoring of raw values where appropriate. We included zero values in cases where data providers indicated that these were legitimate observations. We further identified a single pH value of 1021 which was excluded as it was deemed to be infeasible.

Our default method was to include observations coded as QC 600 (highest quality), 500 (fair), 200 (no quality), or 0 (non-verified). After consultation with MfE, we decided to include observations labelled as QC 200 because this label was associated with most observations collected before the introduction of NEMS QC standards in 2019. We further included QC 0 (virtually no observations) in recognition that the unknown quality of these data is comparable to QC 200. Our default method equated to omitting observations labelled as QC 400 (compromised). The collated dataset contained virtually no observations labelled QC 300 (synthetic) or QC 100 (missing), and these were excluded.

Approximately 2000 observations from lakes in the years 2021 to 2024 (21% of all observations in these years) supplied by Bay of Plenty Regional Council (BOPRC) were labelled as QC 400. Further discussion with BOPRC staff revealed that approximately 80% of observations (from rivers and lakes) were labelled as QC 400 because BOPRC's interpretation of the NEMS 2019 procedures has resulted in an approach that automatically drops the QC code to 400 if samples are not verified as arriving to the lab below 10°C (or less than the sample collection temperature where samples are delivered within two hours of collection), rather than assuming this temperature condition was met. We noted that none of the observations labelled as QC 200 could be verified as arriving to the lab below 10°C from the information available to us. We also noted that our default approach to exclude observations labelled as QC 400 would have greatly reduced the number of sites for which state and trends could be applied within the Bay of Plenty region. We therefore included observations labelled as QC 400 by BOPRC. Observations labelled as QC 400 by all other data providers were excluded.

### Step 4. Monitoring site spatial information.

The following spatial data were associated with each lake monitoring site: LAWA (Land Air Water Aotearoa) identifier, site location, regional council identifier, coordinate, and FENZ lake identifier. Missing monitoring site metadata such as LAWA identifiers or coordinates were identified from previous LAWA water quality datasets and previous environmental reporting datasets where possible. Coordinates were converted to a consistent format and mapped to identify potential errors. FENZ identifiers, where not provided by councils, were identified for each site as the largest lake within 200m, or within 500m if there were no lakes within 200m. FENZ

identifiers were checked by visual inspection of maps. FENZ identifiers were used to link each site to metadata including land cover, depth, and geomorphology class. Where lake depth was not available from FENZ lake depths were identified from alternative sources<sup>3</sup>. Lake depth was used to assign lakes to one of two lake classes: polymictic lakes (<15m) and seasonally stratified and brackish lakes (≥15m). Waiau Lagoon (Southland), which had no FENZ identifier, was classified as shoreline geomorphology class, and its land cover was identified from its position in the digital river network following the methods of Booker et al. (2025).

#### Step 5. Replacing repeat observations.

Repeat observations of the same variable on the same day may have been present in our data despite our intention not to collate frequent observations made for non-SoE purposes. Where there was more than one observation on a day for a given variable at a site, the multiple observations were replaced by the median of the observations.

#### Step 6. Assigning observations to periods.

We assigned each water quality observation metadata describing the date, year, month, bi-month, quarter, and bi-annual periods the observation belonged to. These periods represent the range of monitoring frequencies that exist across all variables and sites in the water quality dataset. A common issue in monthly SoE water quality monitoring arises when sampling is not conducted on the same calendar day each month. For example, if sampling is avoided on weekends or holidays, the timing can shift so that one month ends up without a sample and the next has two. If this happens frequently, the site may not comply with filtering rules (i.e., minimum data requirements for a state or trend assessment to be judged robust – see Section 3.1.3). We identified cases where months with two or more observations were preceded or followed by months missing an observation, and where the time between these two observations is at least two weeks (to ensure independence). Where the previous month has a missing observation, the first observation is assigned metadata associated with the previous month, and where the following month has a missing observation, the last observation is assigned metadata associated with the following month. The observations retain the original date for use in evaluating the trend rate. This data filtering step is a new addition since the last SoE reporting.

---

<sup>3</sup> LAWA, website, NIWA LakeSPI website, Waiau (Te Waewae) Lagoon 2012 Fine Scale Monitoring and Macrophyte Mapping. Prepared For Environment Southland. August 2012, Historical account in Journal of New Zealand Studies NS36 (2023), 55-71, <https://doi.org/10.26686/jnzs.iNS36.8327>

## 3 Analysis methods

### 3.1 Water quality state analyses

#### 3.1.1 Grading of monitoring sites and lakes

Water quality state for lake monitoring sites was graded based on attributes and associated attribute state bands defined by the National Objectives Framework (NOF) of the National Policy Statement – Freshwater Management (NPS-FM; New Zealand Government 2020) (Table 3-1).

Each table of Appendix 2 of the NPS-FM (2020) represents an attribute that must be used to define an objective that provides for a particular environmental value. For example, Appendix 2A, Table 5, defines the ammonia toxicity attribute, which is defined by ammoniacal nitrogen concentrations that will ensure an acceptable level of support for the “Ecosystem health (Water quality)” value. Objectives are defined by one or more numeric attribute states associated with each attribute. For example, for the ammoniacal nitrogen attribute there are two numeric attribute states defined by the annual median and the 95<sup>th</sup> percentile concentrations.

For each attribute, the NOF defines categorical attribute states in four (or five) attribute bands, which are designated A to D (or A to E, in the case of the *E. coli* attribute). The attribute bands represent a graduated range of support for environmental values from high (A band) to low (D or E band). The ranges for attribute states that define each attribute band are defined in Appendix 2 of the NPS-FM (2020). For most attributes, the D band represents a condition that is unacceptable (with the threshold between the C and the D band being referred to as “bottom line”) in any waterbody nationally. In the case of the nitrate (toxicity) and ammonia (toxicity) attributes in the 2020 NPS-FM, the C band is unacceptable, and for the DRP attribute, no bottom line is specified.

The primary aim of the attribute bands designated in the NPS-FM is to provide a basis for objective setting as part of the NOF process. The attribute bands avoid the need to discuss objectives in terms of technically complicated numeric ranges. Each band is associated with a narrative description of the outcomes for values that can be expected if that attribute band is chosen as the objective. However, it is also logical to use attribute bands to provide a grading of the current state of water quality; either as a starting point for objective setting or to track progress toward objectives.

A monitoring site can be graded for each attribute by assigning it to attribute bands (e.g., a site can be assigned to the A band for the nitrate toxicity attribute). The grades are referred to as “NOF grades” in the results below. Site grading is done by using the numeric attribute state (e.g., annual median nitrate-nitrogen) as a compliance statistic. The value of the compliance statistic for a site is calculated from a record of the relevant water quality variable (e.g., the median value is calculated from the observed monthly nitrate-nitrogen concentrations). The site’s compliance statistic is then compared against the numeric ranges associated with each attribute band and a grade assigned for the site (e.g., an annual median total phosphorous concentration of 0.015 mg/L would be graded as “B-band”, because it lies in the range  $>0.01$  and  $\leq 0.02$  mg/L). Note that for attributes with more than one numeric attribute state, we have provided a grade for each numeric attribute state (e.g., for the ammonia (toxicity) attribute, grades are defined for both the median and maximum concentrations).

Table 3-1 provides a summary of the NOF attributes and numeric attribute states calculated as part of this study. We converted units from those shown in Table 2-1 to those shown Table 3-1 when calculating NOF bands. In addition to the NOF attributes in Table 3-1 we also report on lake water quality state for raw (not pH adjusted)  $\text{NH}_4\text{-N}$ , Trophic level index (TLI3 and TLI4), and NNN. For these variables, we report the median of the observations.

Following the methods applied by Whitehead et al. (2021), we used median site NNN concentrations rather than  $\text{NO}_3\text{-N}$  concentrations to grade sites in terms of the NOF nitrate toxicity attribute for the following reasons. The biological mechanisms of nitrite toxicity in freshwater animals are relatively well characterised (Camargo et al. 2005). The primary mechanism is methaemoglobinemia, which occurs when nitrite converts haemoglobin to methaemoglobin. Methaemoglobin cannot bind oxygen, resulting in decreased oxygen transport and tissue hypoxia. Additional mechanisms of nitrite toxicity include electrolyte imbalance and, possibly, the formation of carcinogenic N-nitroso compounds. Nitrite is either ingested directly or generated internally through the bacterial reduction of ingested nitrate. The direct toxic effects of ionic nitrate on freshwater animals have not been identified with certainty. Instead nitrate toxicity is predominantly indirect, mediated by the conversion of nitrate to nitrite by gut bacteria, and the subsequent direct effects of nitrite listed above. Therefore “nitrate toxicity” is more accurately described as “nitrate/nitrite toxicity” (Gehl 2009).

The most common laboratory method for analysing nitrate and nitrite in water samples from NZ freshwater monitoring programmes involves the reduction of all nitrate to nitrite, followed by the colorimetric measurement of nitrite-N concentration based on a standard curve. The results represent the sum of nitrate-N and nitrite-N (abbreviated NNN) in the water samples. Therefore, because “nitrate toxicity” results from the combined ingestion of nitrate and nitrite, and most regional councils monitor NNN concentrations in their freshwater SoE programmes, we used NNN for site grading in lieu of  $\text{NO}_3\text{-N}$ . This approach makes best use of the available nitrate toxicity data and aligns with the approach taken by LAWA and Whitehead et al. (2021).

Along with individual site grades, we also evaluated a lake grade. For lakes with more than one monitoring site, the lake grade is the median of the site grades. For lakes with only site, the lake grade is the same as the site grade. We present tabulated results of the number of lakes and sites assigned to each grade for each NOF attribute. Site NOF grades are mapped for each attribute. NOF grades are summarised in stacked bar charts showing the percentage of lakes assigned each grade, both overall, and within geomorphology and human modified land cover classes.

**Table 3-1: Details of the NOF attributes used to grade the state of lake monitoring sites.**

<b>NPS-FM Reference – NOF Attribute</b>	<b>Calculation guidance</b>	<b>Numeric attribute state description</b>	<b>Units</b>	<b>Abbreviated name</b>
A2A; Table 1 – Phytoplankton (trophic state)		Median concentration of chlorophyll <i>a</i>	mg m <sup>-3</sup>	Chlorophyll <i>a</i> Median
		Maximum concentration of chlorophyll <i>a</i>	mg m <sup>-3</sup>	Chlorophyll <i>a</i> Maximum
A2A; Table 3 – Total nitrogen (trophic state)		Median concentration of total nitrogen	mg m <sup>-3</sup>	TN Median
A2A; Table 4 – Total phosphorous (trophic state)		Median concentration of total phosphorus	mg m <sup>-3</sup>	TP Median
A2A; Table 5 – Ammonia (toxicity)	Based on temperature and pH adjusted Ammoniacal-N	Median concentration of Ammoniacal-N	mg L <sup>-1</sup>	NH <sub>4</sub> -N Median
		95 <sup>th</sup> percentile concentration of Ammoniacal-N	mg L <sup>-1</sup>	NH <sub>4</sub> -N 95 <sup>th</sup> percentile
A2A; Table 9 - <i>Escherichia coli</i>	minimum of 60 samples over a maximum of 5 years	% exceedances over 260 cfu 100 mL <sup>-1</sup>	%	<i>E. coli</i> % >260
		% exceedances over 540 cfu 100 mL <sup>-1</sup>	%	<i>E. coli</i> % >540
		Median concentration of <i>E. coli</i>	cfu 100 mL <sup>-1</sup>	<i>E. coli</i> Median
		95 <sup>th</sup> percentile concentration of <i>E. coli</i>	cfu 100 mL <sup>-1</sup>	<i>E. coli</i> 95 <sup>th</sup> percentile

### 3.1.2 Handling censored values

Following the methods applied by Whitehead et al. (2021), censored values were replaced by imputation for the purposes of calculating the compliance statistics. Left censored values (values below the detection limit(s)) were replaced with imputed values generated using ROS (Regression on Order Statistics; Helsel 2012), following the procedure described in Larned et al. (2015). The ROS procedure produces estimated values for the censored data that are consistent with the distribution of the uncensored values and can accommodate multiple censoring limits. When there are insufficient non-censored data to evaluate a distribution from which to estimate values for the censored observations, censored values are replaced with half of their reported value.

Censored values above the reporting limit were replaced with values estimated using a procedure based on “survival analysis” (Helsel 2012). A parametric distribution is fitted to the uncensored observations and then values for the censored observations are estimated by randomly sampling values larger than the censored values from the distribution. The survival analysis requires a minimum number of observations for the distribution to be fitted; hence in the case that there were fewer than 24 observations, censored values above the detection limit were replaced with the detection limit multiplied by 1.1. The supplementary file outputs provide details about whether and how imputation was conducted for each site by attribute assessment.

### 3.1.3 Time period for assessments and minimum data requirements

When grading sites based on NPS-FM attributes, it is standard practice to define consistent time periods for all sites and to define the acceptable proportion of missing observations (i.e., data gaps) and how these are distributed across sample intervals so that site grades are assessed from comparable data. The time period, acceptable proportion of gaps and representation of sample intervals by observations within the time period are commonly referred to as site inclusion or filtering rules (e.g., Larned et al. 2018) but are also termed “site screening criteria” and “completeness criteria” (Snelder et al. 2021).

The grading assessments were based on a compliance statistic (e.g., the median value of the observations), made for the 5-year time period to end of December 2024. The start and end dates for this period were determined by the availability of quality assured data (see Section 2), reporting time periods defined by MfE, and consideration of statistical precision of the compliance statistics used in the grading of state. The statistical precision of the compliance statistics depends on the variability in the water quality observations and the number of observations. For a given level of variability, the precision of a compliance statistic increases with the number of observations. This is particularly important for sites that are close to a threshold defined by an attribute band because the confidence that the assessment of state is “correct” (i.e., that the site has been correctly graded) increases with the precision of the compliance statistics (and therefore with the number of observations). As a general rule, the rate of increase in the precision of compliance statistics slows for sample sizes greater than 30 (i.e., there are diminishing returns on increasing sample size with respect to precision (and therefore confidence in the assigned grade) above this number of observations; McBride 2005).

In this study, a period of five years represented a reasonable trade-off for grading assessments because it yielded a sample size of 30 or more for many sites and variable combinations. The five-year period for the state analyses is consistent with periods used in previous national water

quality state analyses (e.g., Whitehead et al. 2021). Because water quality data tend to fluctuate seasonally, it is also important that each season is well-represented over the period of record. In New Zealand, it is common to sample either monthly or quarterly, and in these cases, seasons are defined by months or quarters. We therefore applied a rule that restricted site × variable combinations in the state analyses to those with measurements for at least 90% of the sampling intervals in that period (at least 56 of 60 months, 27 of 30 bi-months, or 18 of 20 quarters). Site × variable combinations that did not comply with these rules were excluded from the state analysis. For annual variables TLI3 and TLI4 the nominated minimum sample size requirement was four years.

For grading the *E. coli* attributes, the NPS-FM requires 60 observations over 5 years. For monthly monitoring, this requires collection of all monthly observations (i.e., no missing data). For this study, we relaxed the rule to require observations for 90% of months over the 5-year period (54 observations). Both this relaxation and our default sample number are subjective choices. Therefore, within the supplementary files we provide state assessments for all sites regardless of whether they meet the filtering rules, as well as details about the number of observations and number of years with observations. This will allow MfE to apply tighter or more lenient filtering rules as required.

### 3.1.4 Calculation of percentiles and compliance statistics

For each lake site and variable, we characterised the current state using percentiles (5<sup>th</sup>, 20<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup>, 80<sup>th</sup>, 95<sup>th</sup>) derived from the distribution of measured values for the period 2020 to 2024 (inclusive). All percentiles were calculated using the Hazen method.<sup>4</sup>

Our analysis procedure was designed to account for monitoring records that exhibited irregular sampling frequencies, e.g., multiple observations within a sampling interval (e.g., a month), or changes in sampling frequency over the state period (e.g., from quarterly to monthly monitoring). To avoid a bias in the state assessments towards the period with high sampling frequency, the data was downsampled so that there was only one observation per sampling interval. The sampling interval was selected as the highest frequency sampling interval that would comply with the minimum data requirements. The downsampling procedure involved filtering for the observations closest to the middle of each of the sampling intervals in the state time period.

The compliance statistics specified as “Annual” (maximum, median, 95<sup>th</sup> percentile) in the NPS-FM were calculated from the data pertaining to the entire five-year state period. When reporting state results as box and whisker graphs, the median compliance statistic across sites for each lake was used, so as to not bias the distributions towards lakes with multiple monitoring locations. In the supplementary output files, we have also included the percentiles for data complying with the NEMS guidance, and for all observed data.

### 3.1.5 Relationships between water quality state and catchment land cover

We tallied lakes by each of the evenly spaced categories of human modified land cover. For each water quality variable, we produced box and whisker plots of the median of each lake (median across site medians for lakes with more than one site) in each evenly spaced category

---

<sup>4</sup> (<http://www.mfe.govt.nz/publications/water/microbiological-quality-jun03/hazen-calculator.html>) Note that there are many possible ways to calculate percentiles. The Hazen method produces middle-of-the-road results, whereas the method used in Excel does not (McBride 2005, chapter 8).

of human modified land cover. For *E. coli*, we produced box and whisker plots of the 5<sup>th</sup>, 20<sup>th</sup>, 50<sup>th</sup>, 80<sup>th</sup>, and 95<sup>th</sup> percentiles of each lake in each evenly spaced category of human modified land cover. For *E. coli* we also calculated exceedance measures as the percentage of observations that exceeded 260 and 540 100 mL<sup>-1</sup>, respectively. For each variable, we plotted the median of each lake against human modified land cover. We coloured the points of the scatterplot by dominant land cover (as used by Whitehead et al. 2022) to allow interpretation of which land cover types are contributing to the gradient of human modified land cover. We used linear regressions to relate water quality state to a gradient in human modified land cover. All water quality variable values were log-transformed to improve the normality of residuals within linear regressions.

We also produced box and whisker plots of the percentage of lake observations exceeding 260 cfu 100 mL<sup>-1</sup> and 520 cfu 100 mL<sup>-1</sup> in each evenly spaced category of human modified land cover.

## 3.2 Water quality trend analyses

### 3.2.1 Sampling dates, seasons and time periods for analyses

In trend assessments, there are several reasons why it is important to define the trend period and sampling frequency and to assess whether the observations are adequately distributed over time. First, because variation in many water quality variables is associated with the time of the year, the robustness of trend assessment is likely to be diminished if the observations are biased to particular times of the year. Second, a trend assessment will represent observed patterns from within a predefined time period; essentially that time period demarcated by the first and last observations. The assessment's characterisation of the change in the observations over the time period will not be robust if the observations are not reasonably evenly distributed within the time period. For these reasons, important steps in the data compilation process include specifying the sampling frequency, the time period, and ensuring observations are adequately distributed throughout the time period.

Monitoring programmes are typically designed to sample with a set frequency, (e.g., monthly, quarterly). The trend analysis "time increment" is specified to match this typical sampling frequency (e.g., time increment are months, bi-months or quarters). There is therefore generally an observation for each sample interval (i.e., each time increment within each year). TLI provides an exception to this rule because it is an annual variable calculated for each limnological year.

Two common deviations from the prescribed sampling regime are: (1) a change in sampling interval within the time period; and (2) the collection of more than one observation in a sample interval (e.g., two observations within a month). Both these deviations can occur in the data supplied to us but were mitigated in our analysis after we were able to use information supplied by data providers to identify and remove high-frequency observations that were not undertaken for routine SoE monitoring purposes as described above. Nevertheless, our procedure applied the following methods. For type (1) deviations, we coarsened the sampling interval to define the trend time increment for the part of the record with the higher sampling frequency. This was achieved by taking the observation in the higher frequency part of the record that was closest to the midpoint of the time increments defined by the coarser part of the record. A similar treatment is applied for type (2) deviations, where only the observation closest to the midpoint

of the time increment was retained. The reason for not using a median value is that it will induce a trend in variance, which will invalidate the distributional assumptions of the Mann Kendall S test statistic (Helsel et al. 2020).

The trend at each site was characterised by the rate of change of the central tendency of the observations of each variable through time. Because water quality is constantly varying through time, the evaluated rate of change depends on the period over which the trend is assessed (Ballantine et al. 2010; Larned et al. 2016). Therefore, trend assessments are carried out for specified periods. In the current study, MfE requested that trends be evaluated for periods of 10, 20 and 30 years, ending in December 2024.

MfE also requested trends be evaluated for each of the water quality variables at each monitoring site for rolling windows of 10-years duration starting in 1990 and incrementing by one year (ending 31 December) to a final window ending in 2024 (i.e., a total time period of 34 years). Aggregate rolling trends allowed inspection of stability or transience in proportions of improving trends. The purpose of producing aggregated rolling trends was to investigate temporal patterns with 10-year aggregate trends. We therefore inspected the results of rolling trends to assess the degree to which the proportions of improving trends were either persistent or contained quasi-periodic fluctuations.

For a national study that aims to allow robust comparison of trends between sites and to provide a synoptic assessment of trends across the whole country, such as the present study, it is important that the trends evaluated at each site are commensurate in terms of their statistical power and representativeness of the time period. In these types of studies, it is general practice to define consistent time periods (i.e., trend duration and start date) so that all sites are subjected to the same conditions (i.e., equivalent political, climate, economic conditions). It is also general practice to define the acceptable proportion of gaps and how these are distributed across sample intervals so that the reported trends are assessed from comparable data. The acceptable proportion of gaps and representation of sample intervals by observations within the time period are commonly referred to as site inclusion or filtering rules (e.g., Larned et al. 2018) but this is also termed “site screening criteria” and “completeness criteria”.

Universally applied data requirements or filtering rules have not been developed for trend assessments performed over many sites and variables. The definition of filtering rules is complicated by a trade-off: more restrictive rules, which increase the robustness of the individual trend analyses but will generally exclude numerous sites thereby reducing spatial coverage. In general, this trade-off is also affected by the duration of trend period. Steadily increasing monitoring effort over time means that shorter and more recent trend periods will generally have a larger number of eligible sites.

The application of filtering rules for variables that are measured at quarterly intervals or more frequently requires two steps. First, retain sites for which observations are available for at least X% of the years in the period. Second, retain sites for which observations are available for at least Y% of the sample intervals. For variables that are measured annually such as TLI, the filtering rules are applied by retaining sites for which values are available for at least X% of the years in the trend period.

In this study, for analysis of trend we used filtering rules applied by Larned et al. (2018) and Whitehead et al. (2021), which set X and Y to 80%. Further, the definition of the time increments

specified was flexible to maximise the number of sites that were included. If the site failed to comply with filter rule (2) when time increment was set as months, a coarsening of the data to bi-monthly (once every two months) or quarterly time increments was applied and the filter rule (2) was reassessed. Trends were evaluated based on the highest frequency coarser time increment that complied with filter rule (2). It is noted that the filtering rules imply a tolerance of variable levels of statistical power and temporal representativeness across the sites that were included in the analysis.

### 3.2.2 Seasonality assessment

For many site × variable combinations, observations vary systematically by season (e.g., by month or quarter). In cases where seasons are a major source in variability, accounting for the systematic seasonal variation should increase the statistical power of the trend assessment (i.e., increase the confidence in the estimate of direction and rate of the trend). The purpose of a seasonality assessment was to identify whether seasons explain variation in the water quality variable. If this was true, then seasonal versions of the trend assessment procedures were used at the trend assessment step (Section 3.2.3).

In previous national scale assessments, the season assigned to a site × variable combination was limited to the sampling interval. However, for our trend assessment, season can be defined separately from the time increment of the assessment. The possible options for season are the selected time increment, and any other increments (i.e., monthly, bi-monthly, quarterly, biannually) that are whole multiples of the time increment. We evaluated seasonality for each potential season using the Kruskal-Wallis multi-sample test for identical populations. This is a non-parametric ANOVA that determines the extent to which season explains variation in the water quality observations. Following Hirsch et al. (1982), we identified season × site × variable combinations as being seasonal based on the  $p$ -value from the Kruskal-Wallis test with  $\alpha=0.05$ . At this point, a check was also performed to determine whether each season option would have sufficient data variability within each season to undertake the trend assessment ( $\geq 5$  non-censored values and/or  $\geq 3$  unique non-censored values) (Section 3.2.5). The selected season is the time increment that meets minimum variability requirements and has a  $p$  value  $< 0.05$  or, in the case that more than season increment meets these thresholds, has the largest Kruskal-Wallis test statistic. For these sites × variable combinations, subsequent trend assessments followed the “seasonal” variants, described in Section 3.2.3.

Our approach varies slightly from that used in the previous national assessment, where season and time increment were required to be the same interval, and if a seasonal assessment did not meet minimum within season variability requirements, both the time increment and season were coarsened, and the seasonality test repeated. The approach used in this study reduces unnecessary loss of data in the coarsening process.

The choice of  $\alpha$  is subjective and a value of 0.05 is associated with a very high level of certainty (95%) that the data exhibit a seasonal pattern. In our experience there are generally diminishing differences between the seasonal and non-seasonal trend assessments associated with the Kruskal-Wallis test for  $p$ -values values larger than 0.05 (Helsel et al. 2020).

### 3.2.3 Analysis of trends

The purpose of trend assessment is to evaluate the direction (i.e., improving or worsening) and rate of the change in the central tendency of the observed water quality values over the period of

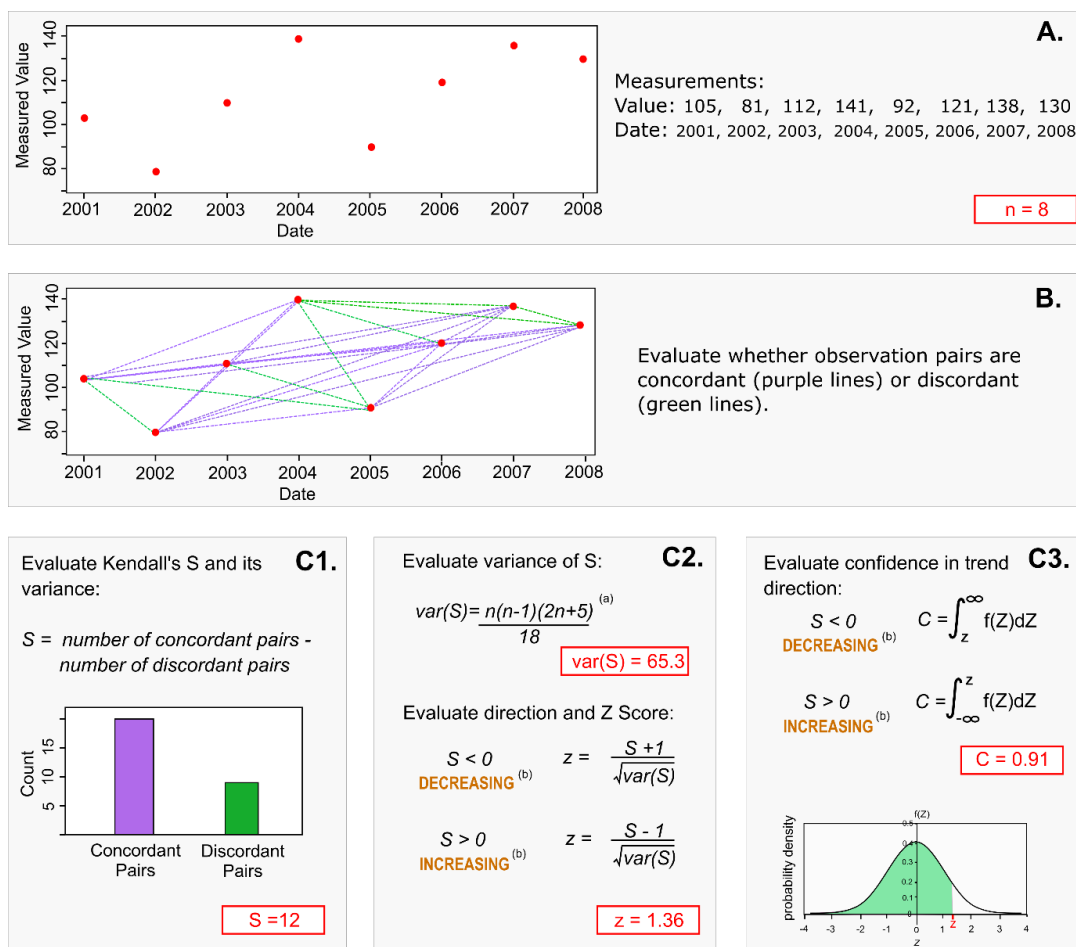
analysis (i.e., the trend). Because the analysed data represent observations of the water quality over the period of analysis, there is uncertainty about the conclusions drawn from their analysis. Therefore, statistical models are used to determine the direction and rate of the trend and to evaluate the uncertainty of these determinations.

We evaluated trends using the LWPTrends functions (Fraser and Snelder 2025) that are implemented in the R statistical computing software (R Core Team 2023). A brief description of the theoretical basis for these functions is described below.

### Assessments of trend directions

The trend direction and the confidence in trend direction were evaluated using either the Mann Kendall assessment or the Seasonal Kendall assessment. Although the non-parametric Sen slope regression also provides information about trend direction and its confidence, the Mann Kendall assessment is recommended, rather than Sen slope regression, because the former more robustly handles censored values. However, Sen slope regression is the recommended method for assessing the trend rate.

The Mann Kendall assessment requires no a priori assumptions about the distribution of the data but does require that the observations are randomly sampled and independent (no serial correlation) and that there is a sample size of  $\geq 8$ . Both the Mann Kendall and Seasonal Kendall assessments are based on calculating the Kendall S statistic, which is explained diagrammatically in Figure 3-1.



**Figure 3-1: Schematic diagram demonstrating how the Kendall S statistic and confidence in trend direction (C) is calculated.** Notes: [a] the calculation of the variance in S has some adjustments to account for ties (numerically equal values, or ties in time) and censored values. Details of these adjustments can be found in (Gilbert 1987; Helsel 2005, 2012). [b] There is a third alternative, where  $S=0$ . In this case C is 0.5, and the trend direction is classified as “indeterminate”. Values of S equal to -1 or 1 will also result in a Z value of 0, and a C value of 0.5 and the trend direction is similarly classified as “indeterminate”.

The Kendall S statistic is calculated by first evaluating the differences between all pairs of water quality observations (Figure 3-1 A and B). Positive differences are termed “concordant” (i.e., the observations increase with increasing time) and negative differences are termed “discordant” (i.e., the observations decrease with increasing time). Pairs of observations that are tied in value or tied in time<sup>5</sup> (i.e., have the same time increment-Year) are assigned differences of zero (Gilbert, 1987, p. 243). The Kendall S statistic is the number of concordant pairs minus the number of discordant pairs (Figure 3-1, C1). The water quality trend direction is indicated by the sign of S with a positive or negative sign indicating an increasing or decreasing trend, respectively (Figure 3-1, C2).

<sup>5</sup> Accounting for ties in time is an addition to the analysis since Whitehead et al. (2021). This addition is to accommodate seasonal analyses where the Season time increment is larger than the data TimeIncr (time increment).

The seasonal version of the Kendall S statistic  $S$  is calculated in two steps. First, for each season, the S statistic is calculated in the same manner as shown in Figure 3-1 but for data pertaining to observations in each individual season. Second,  $S$  is the sum of values over all seasons ( $S = \sum_1^n S_i$ ), where  $S_i$  is the number of concordant pairs minus the number of discordant pairs in the  $i^{th}$  season and  $n$  is the number of seasons. The variance of  $S$  is calculated for each season and then summed over all seasons.

The sign (i.e., + or -) of the S statistic calculated from the sample represents the best estimate of the population trend direction but is uncertain (i.e., the direction of the population trend cannot be known with certainty). A continuous measure of confidence in the assessed trend direction can be determined based on the posterior probability distribution of  $S$ , the true (i.e., population) difference in concordant and discordant pairs (Snelder et al. 2022). The posterior probability distribution of  $S$  is given by a normal distribution with mean of  $S$  and variance of  $var(S)$ . The confidence in assessed trend direction can be evaluated as the proportion of the posterior probability distribution that has the same sign as  $S$ .

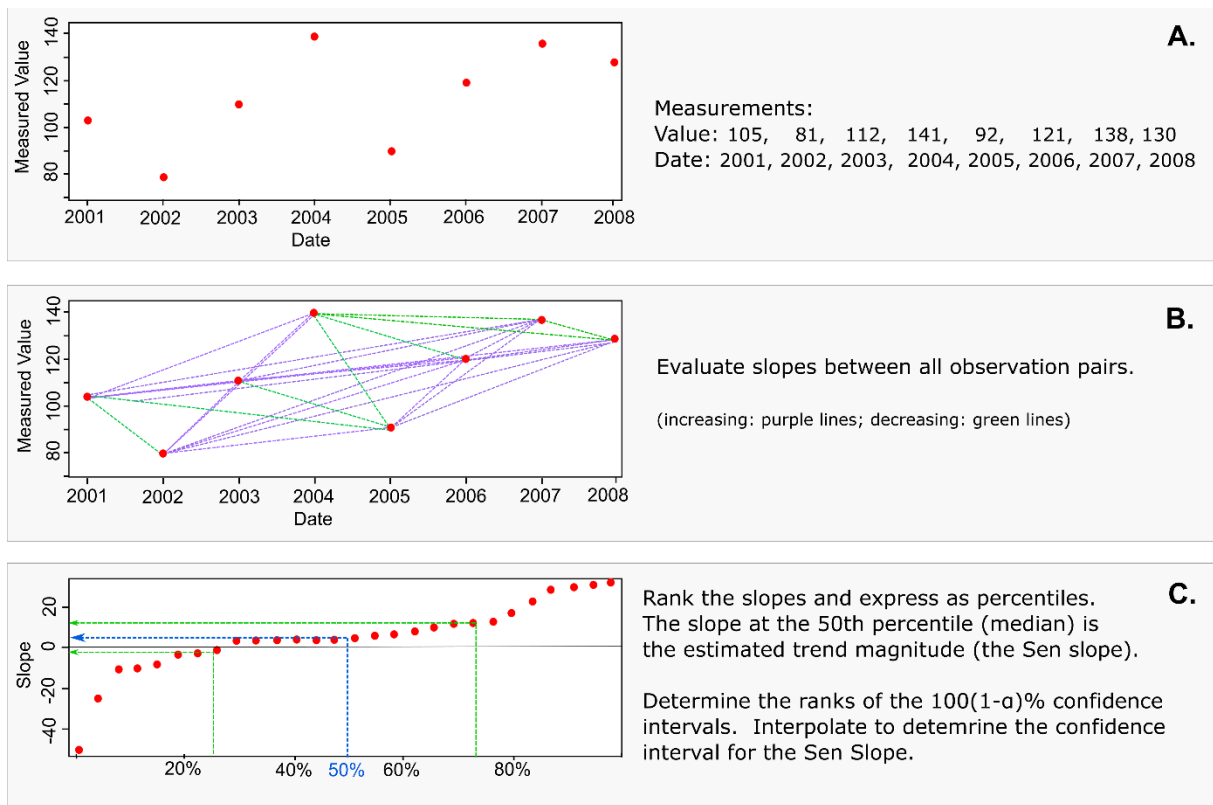
In practice, confidence can be calculated by first transforming the value of  $S = 0$  on the posterior probability distribution into a standard normal deviate,  $Z$  (panel C2).  $C$  is then calculated as area under the standard normal distribution to the left ( $Z > 0$ ) or right ( $Z < 0$ ) of the value of  $Z$ , using the quantile function for the normal distribution.

The value  $C$  can be interpreted as the probability that the sign of the calculated value of  $S$  indicates the direction of the population trend (i.e., that the calculated trend direction is correct). The value  $C$  ranges between 0.5, indicating the sign of  $S$  is equally likely to be in the opposite direction to that indicated by the true trend, to 1, indicating complete confidence that the sign of  $S$  is the same as the true trend.

As the size of the sample (i.e., the number of observations) increases, confidence in the trend direction increases. When the sample size is very large,  $C$  can be high, even if the trend rate is very low. It is important therefore that  $C$  is interpreted correctly as the confidence in direction and not as the importance of the trend. As stated at the beginning of this section; both trend direction and trend rate are relevant and important aspects of a trend assessment.

### Assessments of trend rates

The method used to assess trend rates is based on non-parametric Sen slope regressions of water quality observations against time. The Sen slope estimator (SSE; Hirsch et al. 1982) is the slope parameter of a non-parametric regression. SSE is calculated as the median of all possible inter-observation slopes (i.e., the difference in the measured observations divided by the time between sample dates) (Figure 3-2).



**Figure 3-2: Schematic diagram of the calculation of the Sen slope, which is used to characterise trend rate.** (from Snelder et al. 2021).

The seasonal version of the SSE is used in situations where there are significant (e.g.,  $p \leq 0.05$ , as evaluated using a Kruskal-Wallis test) differences in water quality measurements between “seasons”. Seasons are defined primarily by sampling intervals, which are commonly monthly or quarterly for water quality monitoring but can also be defined as whole multiples of the sampling interval (e.g., monthly sampling interval and quarterly season). The Seasonal Sen Slope Estimator (SSSE) is calculated in two steps. First, for each season, the median of all possible inter-observation slopes is calculated in same manner as shown in but for data pertaining to observations in each individual season. When the time increment is smaller than the season time increment (i.e., there is generally more than one observation per season by year) the inter-observation slopes are excluded for pairs of observations that are in the same season-year (they are treated as ties in time), but all slopes in the season of one year are compared with all slopes in the season of the other years (Gilbert 1987, p. 218). Second, SSSE is the median of the seasonal values. The relative Sen Slope Estimator (RSSE) is the Sen Slope divided by the median value from the observation data and expresses the trend rate as a percentage change per year.

Uncertainty in the assessed trend rate is evaluated following a methodology outlined in Helsel et al. (2020). To calculate the 100(1- $\alpha$ )% two-sided symmetrical confidence interval about the fitted slope parameter, the ranks of the upper and lower confidence limits are determined, and the slopes associated with these observations are applied as the confidence intervals.

### 3.2.4 Handling censored values

Censored values are managed differently to non-censored values by the non-parametric trend assessment methods described in Section 3.2.3. It is therefore important that censored values are correctly identified in the data. Detection limits or reporting limits that have changed through the trend period (often due to analytical changes) can induce trends that are associated with the changing precision of the measurements rather than actual changes in the variable. This possibility needs to be accounted for in the trend analysis, and this is another reason that it is important that censored values are correctly identified in the data.

Censored values in the data used to calculate Kendall's S and its p-value are robustly handled in the manner recommended by Helsel (2005, 2012). Briefly, for left-censored data, increases and decreases in a water quality variable are identified whenever possible. Thus, a change from a censored data entry of <1 to a measured value of 10 is considered an increase. A change from a censored data entry of <1 to a measured value 0.5 was considered a tie, as is a change from <1 to a <5, because neither can definitively be called an increase or decrease. Similar logic applied to right censored values. The information about ties is used in the calculation of the Kendall S statistic and its variance following Helsel (2012) and this provides for robust calculation of the p-value associated with the Kendall test. This approach is implemented in LWPTrends using source code from the NADA package. The method is robust to changes in detection limit over time.

Note that as the proportion of censored values increases, the proportion of ties increases and the confidence in the trend direction decreases. Therefore, confidence in direction tends to be low when trends are calculated from data with high proportions of censored observations.

When calculating Sen slopes, the inter-observation slope cannot be definitively calculated between any combination of observations in which either one or both are censored. Therefore, when SSE and SSSE (i.e., Sen slopes) are calculated by the LWPTrends library, the censored data entries are substituted with their corresponding raw values (i.e., the numeric component of a censored data entry) multiplied by a factor (0.5 for left-censored and 1.1 for right-censored values). This ensures that any measured value that is equal to a raw value is treated as being larger than the censored value if it is left-censored value and smaller than the censored value if it is right-censored. The inter-observation slopes associated with the censored values are therefore imprecise (because they are calculated from the substituted values). However, because the Sen slope is the median of all the inter-observation slopes, the Sen slope is unaffected by censoring when a small proportion of observations are censored. As the proportion of censored values increase, the probability that the Sen slope is affected by censoring increases. The outputs from the trend assessment provide an "analysis note" to identify Sen Slopes where one or both of the observations associated with the median inter-observation slope is censored.

In previous versions of LWPTrends used to evaluate national water quality trends (Whitehead et al. 2021), the influence of changes in detection limits over time on the Sen slope were handled through a high censor filter. Briefly, this filter replaced all observations below the highest detection limit (or some user specified value) to the highest detection limit and marked these observations as censored. We now implement a different approach to handle changes in detection limit when calculating trend rates that follows the logic used in the calculation of Kendall's S. Inter-observation slopes are considered to be ties and set to zero, regardless of

their values, when: (1) both observations were either left or right censored, (2) when one observation is left censored and larger than the other non-censored observation; (3) when one observation is right censored and smaller than the other non-censored observation. This approach has the advantage that the user decisions around it and how to implement a high censor filter are removed.

**3.2.5 Interpretation of trends**

The trend assessment procedures used here allow a more nuanced inference than the categorical measure of confidence associated with Null Hypothesis Significance Testing (i.e., significant or not significant). The confidence in direction (C) can be transformed into a continuous scale of confidence the trend was decreasing ( $C_d$ ). For all trends with  $S < 0$ ,  $C_d = C$ , and for all  $S > 0$  a transformation is applied so that  $C_d = 1 - C$ .  $C_d$  ranges from 0 to 1.0. When  $C_d$  is very small, a decreasing trend is highly unlikely, which because the outcomes are binary, is the same as an increasing trend is highly likely.

To assist with interpretation of confidence in trend direction, trends are categorised based on the confidence that a trend was improving. This involved firstly converting  $C_d$  into a confidence that a trend was improving ( $C_i$ ). Improvement is indicated by decreasing trends for all the water quality variables in this study ( $C_i = C_d$ ) except for Secchi (for which increasing trends indicate improvement). For these variables,  $C_i$  is the complement of  $C_d$  (i.e.,  $C_i = 1 - C_d$ ).

The trend for each site x variable combination was assigned a categorical level of confidence that the trend was improving according to its evaluated confidence, direction and the categories shown in Table 3-2.

**Table 3-2: Level of confidence categories used to convey the confidence that the trend direction was improving.**

<b>Categorical level of confidence trend was improving</b>	<b>Value of <math>C_i</math> (%)</b>
Very likely improving	$0.90 < C_i \leq 1.00$
Likely improving	$0.67 < C_i \leq 0.90$
Low confidence	$0.33 < C_i \leq 0.67$
Likely worsening	$0.10 < C_i \leq 0.33$
Very likely worsening	$0.00 \leq C_i \leq 0.10$

Some trends were classified as “not analysed” for either of two reasons:

1. When a large proportion of the values were censored (data has <5 non-censored values and/or <3 unique non-censored values). This arises because trend analysis is based on examining differences in the value of the variable under consideration between all pairs of sample occasions. When a value is censored, it cannot be compared with any other value and the comparison is treated as a “tie” (i.e., there is no change in the variable between the two sample occasions). When there are many ties there is little information content in the data and a meaningful statistic cannot be calculated.
2. When there is no, or very little, variation in the data, which also results in ties. This can occur because laboratory analysis of some variables has low precision (i.e.,

values have few or no significant figures). In this case, many samples have the same value, and this then results in ties.

### 3.2.6 Aggregation of lake trends

Aggregating water quality trend results from multiple lakes is intended to indicate water quality changes over a domain of interest (e.g., environmental classes, regions, national). In the present study, we aggregated trend results using both trend magnitudes and confidence in trend direction. In to avoid bias towards lakes with multiple sites, lake trend magnitudes and confidence in trend directions were summarised by taking the median across multiple sites within a lake.

The distributions of trend magnitude across lakes were characterised using box and whisker plots of the relative Sen slope estimates (RSSE) and relative seasonal Sen slope estimates (RSSSE). Sen slopes were relativised by dividing the SSE and SSSE values by the duration of the trend period to give estimates of temporal change in % yr<sup>-1</sup>.

We used two different approaches for aggregating trend directions. The first approach involved the calculation of the aggregate proportion of lakes in each categorical level of confidence that the trend was improving (shown in Table 3-2) for each variable; these values were plotted as colour coded stacked bar charts. These charts provide a graphical representation of the proportions of increasing and improving trends at the levels of confidence indicated by the categories. We also used this approach for each of the outputs of the 10-year trends for rolling windows. Results for rolling time windows were only shown where there were at least 20 lakes within the 10-year time window. This was an arbitrary cut-off point selected to minimise bias that might be associated with a small sample size but maximise the number of time windows that were reported.

The second approach also utilises the confidence that the true trend was improving to provide a probabilistic estimate of the proportion of improving lake-specific trends ( $P_i$ ) within a geographic or environmental domain. For a given water quality variable, the trends at multiple monitoring sites distributed across a domain of interest can be assumed to represent independent samples of the population of trends, for all sites within that domain.

The statistic  $P_i$  is calculated by letting the sampled lakes within this domain be indexed by  $m$ , so that  $m \in \{1, \dots, M\}$ , and letting  $I$  be a random Bernoulli distributed variable which takes the value 1 with probability  $p = C_i$  and the value 0 with probability  $q = 1 - C_i$  (where  $C_i$  is the confidence that the trend was improving, as described in Section 3.2.3). Therefore,  $I_m = 1$  denotes an improving trend at lake  $s \in \{1, \dots, M\}$  when the estimated  $p_m \geq 0.5$  and a worsening trend as 0 when  $p_m < 0.5$ . Then, the estimated proportion of sites with improving trends in the domain is:

$$P_i = \sum_{m=1}^M I_m / M$$

Because the variance of a random Bernoulli distributed variable is  $Var(I) = p(1 - p)$ , and assuming the lake trends are independent, the estimated variance of  $P_i$  is:

$$\text{Var}(P_i) = \frac{1}{M^2} \sum_{m=1}^M \text{Var}(I_m) = \frac{1}{M^2} \sum_{m=1}^M p_m(1 - p_m)$$

$P_i$  and its variance represent an estimate of the population proportion of improving trends, within a spatial or environmental domain, and the uncertainty of that estimate. It is noted that the proportion of worsening trends is the complement of the result (i.e.,  $1 - P_i$ ).

If there is spatial correlation in the assessed trend directions at the individual sites, the assumption of independence, when calculating  $P_i$  and its variance as above, is violated (Douglas et al. 2000). Spatial correlation means that the effective sample size of the dataset is less than the number of sites and this results in under-estimation of the variance and over-estimation of confidence in the  $P_i$  statistic. The method of Douglas et al. (2000) can be used to calculate an adjusted version of the variance of  $P_i$  as:

$$\text{Var}(P_i) = \frac{1}{M^2} \left[ \sum_{k=1}^M \text{Var}(I_k) + 2 \sum_{k=1}^{M-1} \sum_{l=1}^{M-k} \text{Cov}(I_k, I_{k+l}) \right]$$

where the covariance between monitoring sites is calculated as:

$$\text{Cov}(I_k, I_{k+l}) = \sqrt{\text{Var}(I_k)\text{Var}(I_{k+l})\rho_{k,k+l}^c}$$

and where  $\rho_{k,k+l}^c$  is replaced by the sample cross-correlation coefficient  $r_{k,k+l}$  (Yue and Wang 2002), which is computed from the observation time series at lake  $k$  and  $k+l$  as:

$$r_{k,k+l} = \frac{\frac{1}{N} \sum_{i=1}^N (x_{k,i} - \bar{x}_k)(x_{k+l,i} - \bar{x}_{k+l})}{\sqrt{\text{Var}(x_k)\text{Var}(x_{k+l})}}$$

where  $x_k^i$  and  $x_{k+l}^i$  represent one set of  $N$  concurrent observations at the sites.

The estimated variance of  $P_i$  can be used to construct 95% confidence intervals<sup>6</sup> around the  $P_i$  statistics as follows:

$$CI_{95} = P_i \pm 1.96 \times \sqrt{\text{Var}(P_i)}$$

We calculated  $P_i$  and its confidence interval for all water quality variables and for domains of interest defined by the entire country, and by the four human modified land cover classes defined in Section 2.6.

---

<sup>6</sup> Note that +/- 1.96 are approximately the 2.5<sup>th</sup> and 97.5<sup>th</sup> percentile of a standard normal distribution.

## 4 Results – lake state

### 4.1 State summary statistics

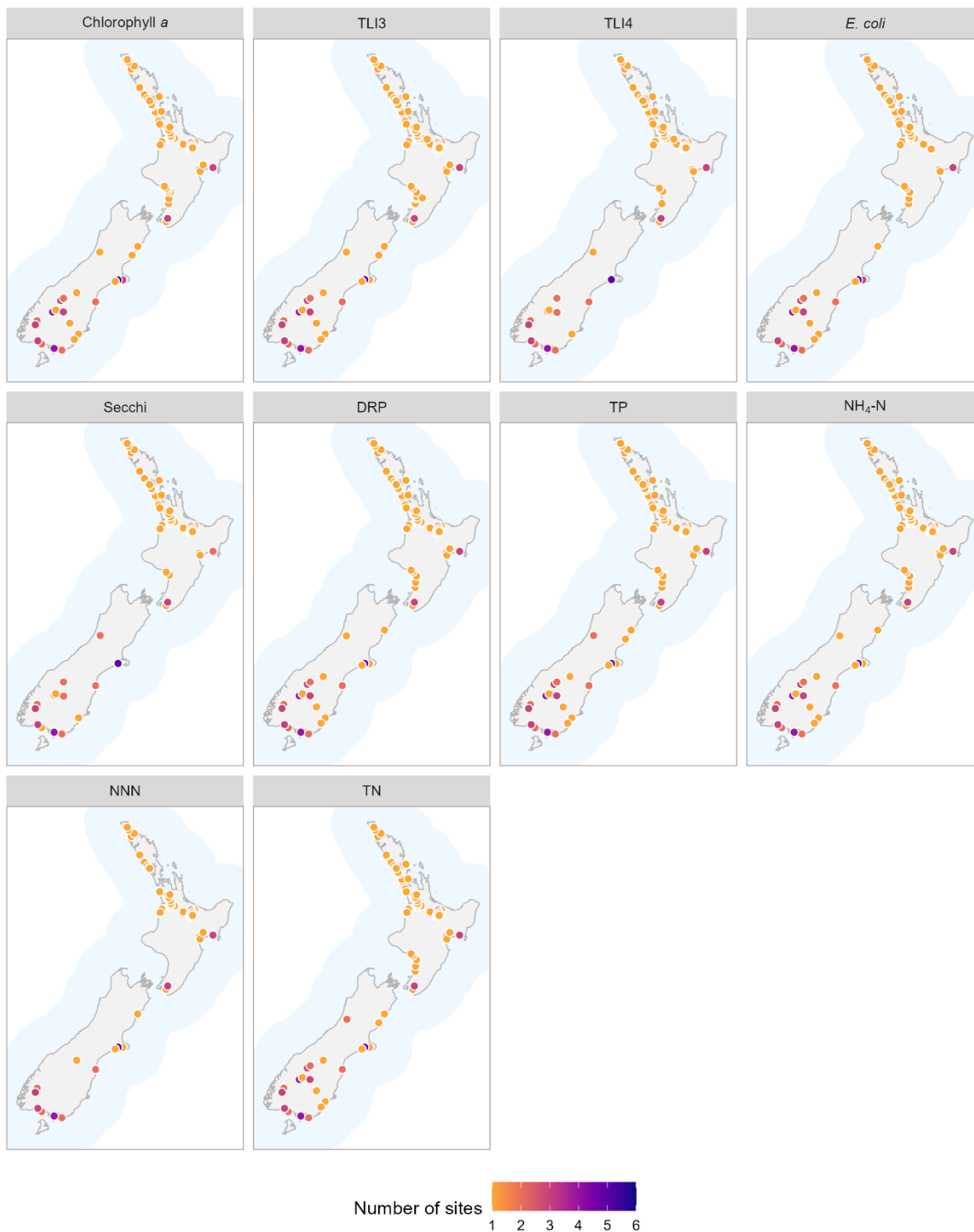
Between 79 and 123 lakes (including 102 to 154 sites) met the filtering rules for the state analysis of nutrients, Secchi, chlorophyll *a*, and TLI; the number of qualifying lakes varied by water quality variable and by both geomorphology class (Table 4-1) and human modified land cover class (Table 4-2). The number of lake × variable combinations is higher than in the dataset used in the previous analysis of lake state (Whitehead et al. 2021). The geographic distribution of lakes in the state dataset is shown in Figure 4-1. The qualifying lakes are sparsely and unevenly distributed across both the North and South Islands. Only a small proportion of observations were censored for many site × variable combinations (Figure 4-2). A large proportion of observations were censored at a small number of sites for some variables (chlorophyll *a*, Secchi, TN, TP), while others (DRP, NH<sub>4</sub>-N, *E. coli*, NNN) exhibited a larger proportion of censored observations at more sites compared to the other variables. The median observation was below the detection limit for 25% of sites for NH<sub>4</sub>-N, 26% of sites for DRP, and 10% of sites for *E. coli* and NNN. The proportion of data with QC information has increased at a rapid rate since introduction of QC standards in 2019, although the majority of recent observations were not associated with QC information (Figure 4-3).

**Table 4-1: Number of lakes (and sites, in brackets) by geomorphology class and water quality variable included in the state analyses.** Note NH<sub>4</sub>-N (Adj.) is the adjusted ammoniacal-N.

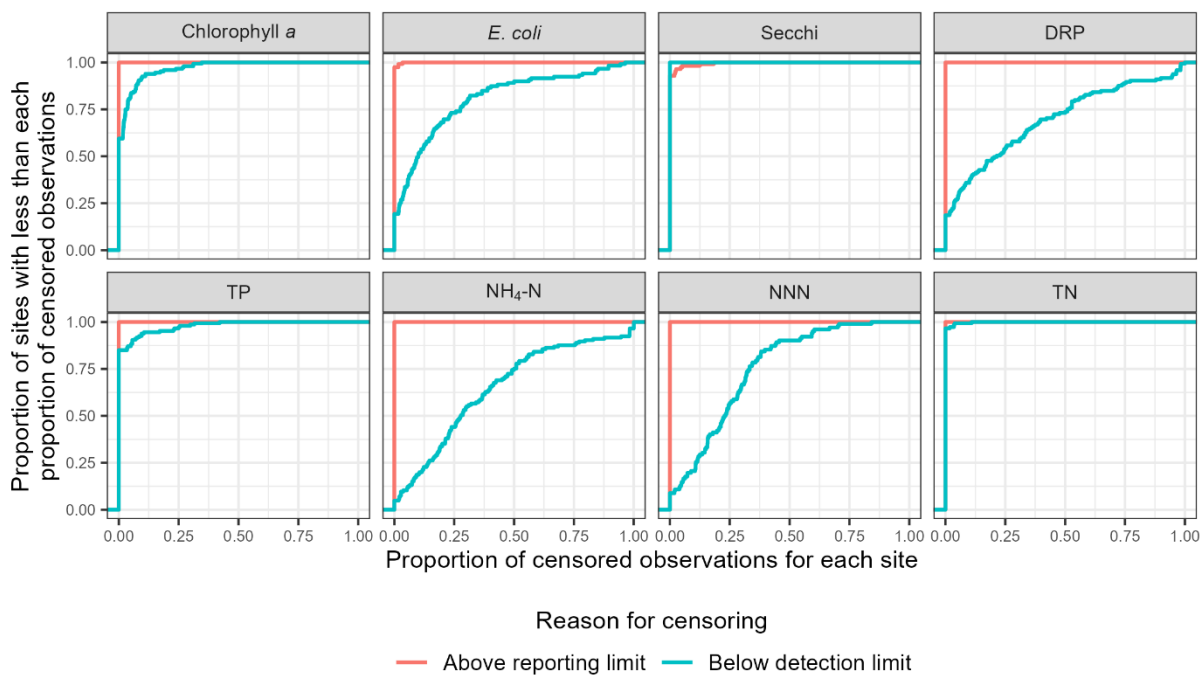
Variable	Number of lakes (and sites)								
	Total	Glacial	Dam	Volcanic	Aeolian	Landslide	Riverine	Peat	Shoreline
Chlorophyll <i>a</i>	114 (148)	8 (17)	4 (7)	17 (20)	47 (48)	3	14 (16)	13	8 (24)
TLI3	123 (154)	8 (17)	4 (7)	17 (20)	51 (52)	3	16 (18)	16	8 (21)
TLI4	99 (124)	6 (10)	2 (4)	14 (17)	38 (39)	3	13 (15)	16	7 (20)
<i>E. coli</i>	94 (119)	5 (11)	3 (6)	3	48 (49)	2	13	13	7 (22)
Secchi	89 (112)	6 (11)	2 (4)	14 (17)	32	3	12 (14)	13	7 (18)
DRP	114 (145)	8 (17)	3 (6)	17 (20)	48 (49)	3	14 (16)	13	8 (21)
TP	115 (147)	8 (18)	4 (7)	17 (20)	48 (49)	3	14 (16)	13	8 (21)
NH <sub>4</sub> -N (adj.)	113 (144)	7 (16)	3 (6)	17 (20)	48 (49)	3	14 (16)	13	8 (21)
NH <sub>4</sub> -N	114 (145)	8 (17)	3 (6)	17 (20)	48 (49)	3	14 (16)	13	8 (21)
NNN	79 (102)	3 (6)	1 (2)	16 (19)	23 (24)	3	12 (14)	13	8 (21)
TN	115 (147)	8 (18)	4 (7)	17 (20)	48 (49)	3	14 (16)	13	8 (21)

**Table 4-2: Number of lakes (and sites, in brackets) by percentage of catchment human modified land cover and water quality variable included in the state analyses.** Note NH<sub>4</sub>-N (Adj.) is the adjusted ammoniacal-N.

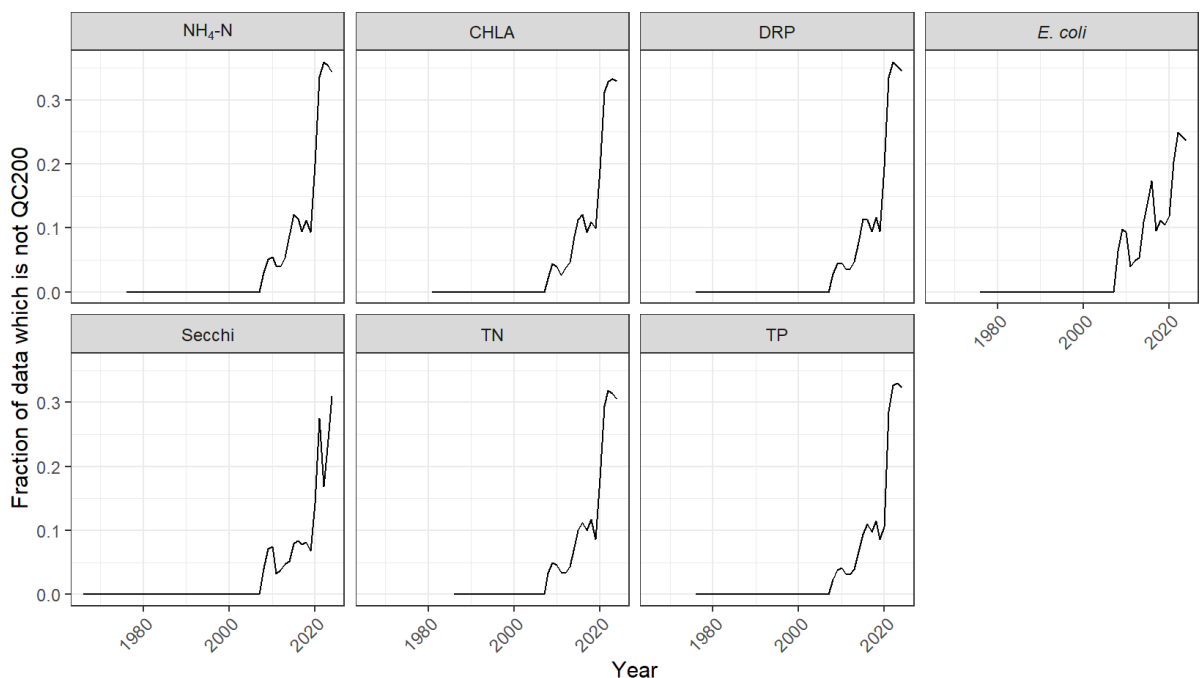
Variable	Number of lakes (and sites)				
	Total	0–25%	25–50%	50–75%	75–100%
Chlorophyll <i>a</i>	114 (148)	14 (27)	19 (22)	33 (45)	48 (54)
TLI3	123 (154)	14 (27)	19 (22)	34 (44)	56 (61)
TLI4	99 (124)	11 (18)	18 (21)	29 (39)	41 (46)
<i>E. coli</i>	94 (119)	9 (19)	13 (14)	25 (34)	47 (52)
Secchi	89 (112)	11 (19)	17 (19)	26 (34)	35 (40)
DRP	114 (145)	14 (27)	19 (22)	33 (43)	48 (53)
TP	115 (147)	14 (28)	19 (22)	33 (43)	49 (54)
NH <sub>4</sub> -N (adj.)	113 (144)	13 (26)	19 (22)	33 (43)	48 (53)
NH <sub>4</sub> -N	114 (145)	14 (27)	19 (22)	33 (43)	48 (53)
NNN	79 (102)	5 (10)	15 (18)	28 (38)	31 (36)
TN	115 (147)	14 (28)	19 (22)	33 (43)	49 (54)



**Figure 4-1: Locations of monitored lakes and number of sites per lake used for state analyses of lake water quality.**



**Figure 4-2: Proportion of censored data for each site × variable combination for which state was assessed.** Note TLI3 and TLI4 are not included as the aggregate annual values are not censored.



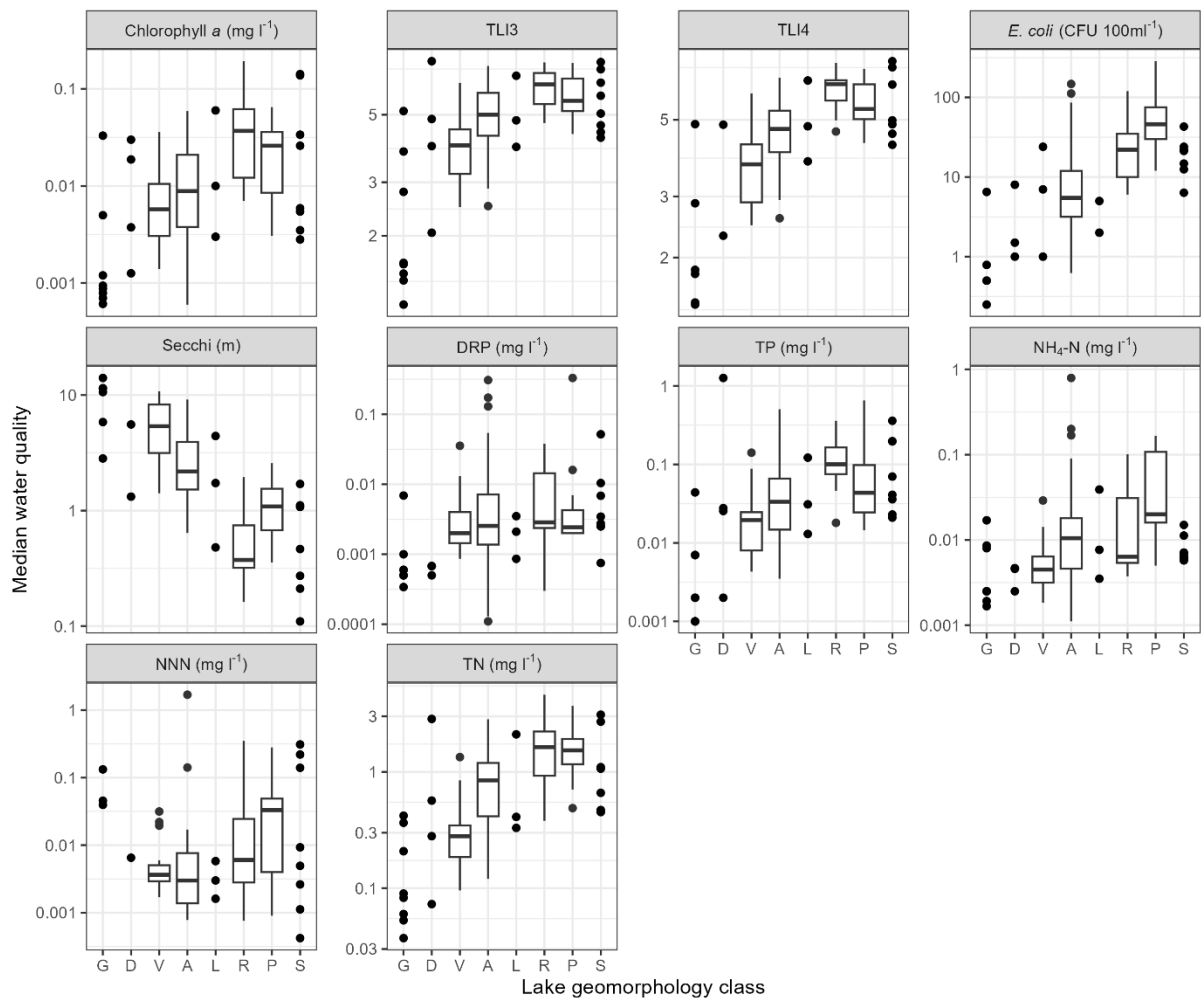
**Figure 4-3: For each variable, the proportion of analysed data that was not labelled QC 200.** The remaining proportion was either labelled as QC 200 (i.e., no quality) by the data provider or had missing QC information.

The distributions of lake-medians for the lake water quality variables for the 2020–2024 period are summarized as box-and-whisker plots by lake geomorphology class in Figure 4-4 and by lake

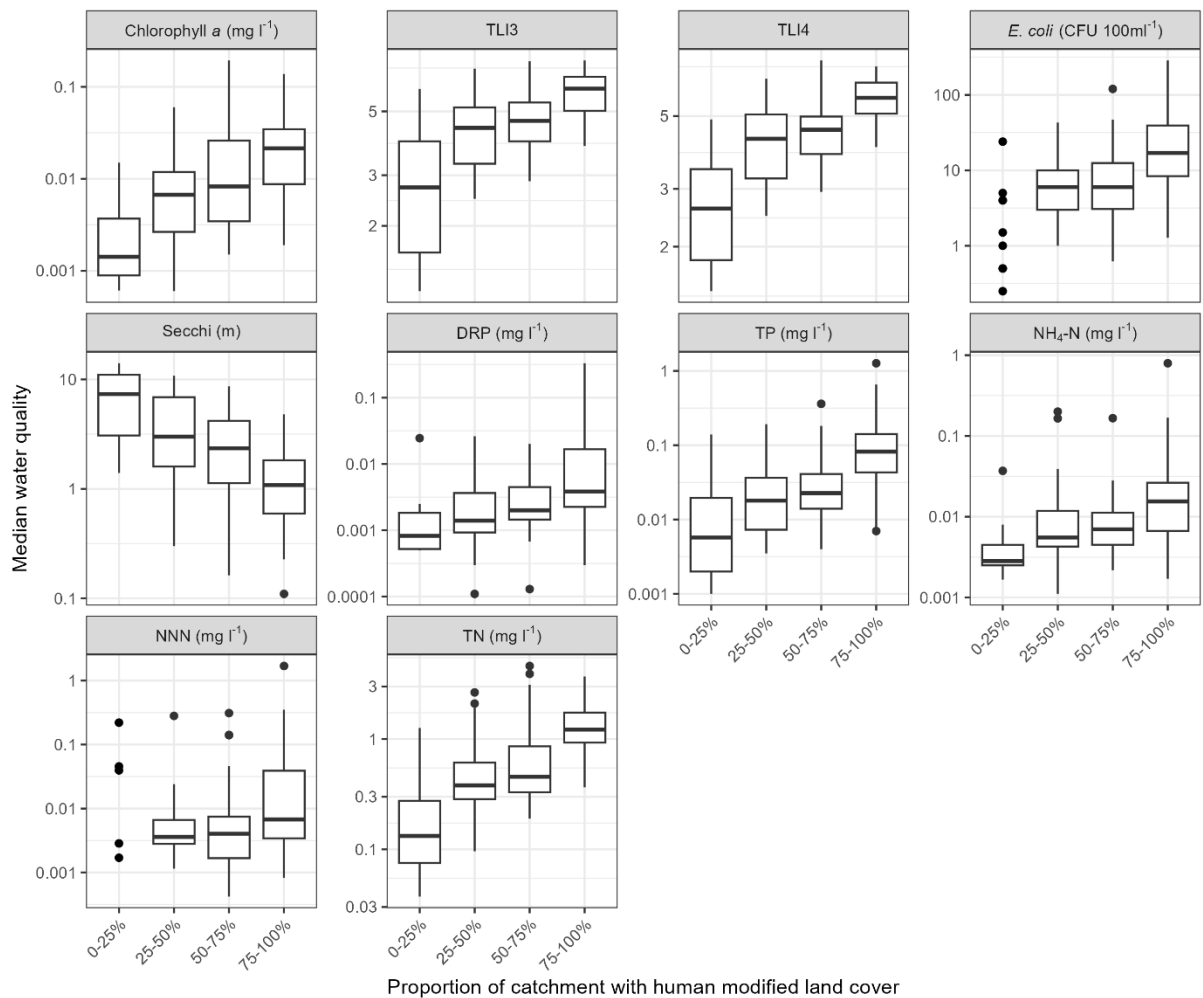
human modified land cover class in Figure 4-5. These plots indicate that some variability in the most widely measured variables is associated with both lake geomorphology and human modified land cover class. All variables show increasing median values (or decreasing for Secchi) with increasing proportion of upstream human modified land cover. Between-geomorphology class differences in lake median values were greatest for TLI, *E. coli*, Secchi, and TP. General patterns between lake geomorphology class and water quality variables medians were present. Variables were generally higher (or lower for Secchi) within glacial (G), dam (D), volcanic (V), aeolian (A), landslide (L) classes, with the riverine (R), peat (P) and shoreline (S) classes consistently having the highest (or lowest for Secchi) median values.

The distribution of *E. coli* concentration percentiles (5<sup>th</sup>, 20<sup>th</sup>, 50<sup>th</sup>, 80<sup>th</sup>, and 95<sup>th</sup>) are shown by geomorphology class in Figure 4-6 and by human modified land cover class in Figure 4-7. The various percentiles of *E. coli* concentrations were generally highest in the riverine and peat geomorphology classes and were generally lower in the glacial, dam, volcanic and aeolian classes (Figure 4-6). Medians of *E. coli* concentration percentiles generally increased with increasing proportions of catchment human modified land cover (Figure 4-7). However, there was overlap in the distributions between catchment human modified land cover classes.

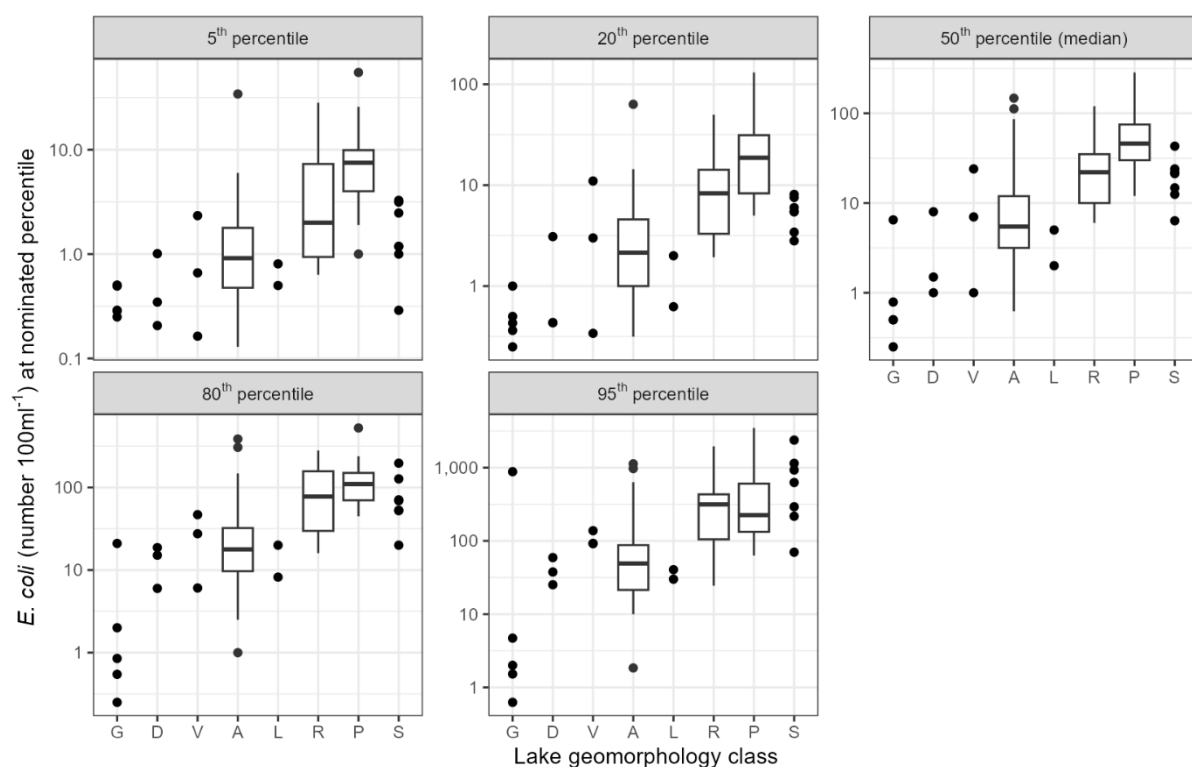
The distribution of the two *E. coli* exceedance measures (the percentage of observations that exceeded 260 and 540 cfu 100 mL<sup>-1</sup>) are shown by geomorphology class in Figure 4-8 and by human modified land cover class in Figure 4-9. Strong patterns between *E. coli* exceedances and geomorphology class were not evident (Figure 4-8). Medians of *E. coli* exceedances generally increased with increasing proportions of catchment human modified land cover (Figure 4-9). The median of each *E. coli* percentile class was highest in the highest human modified land cover class (75–100%) (Figure 4-9). However, there was considerable overlap in the various *E. coli* metrics between human modified land cover classes.



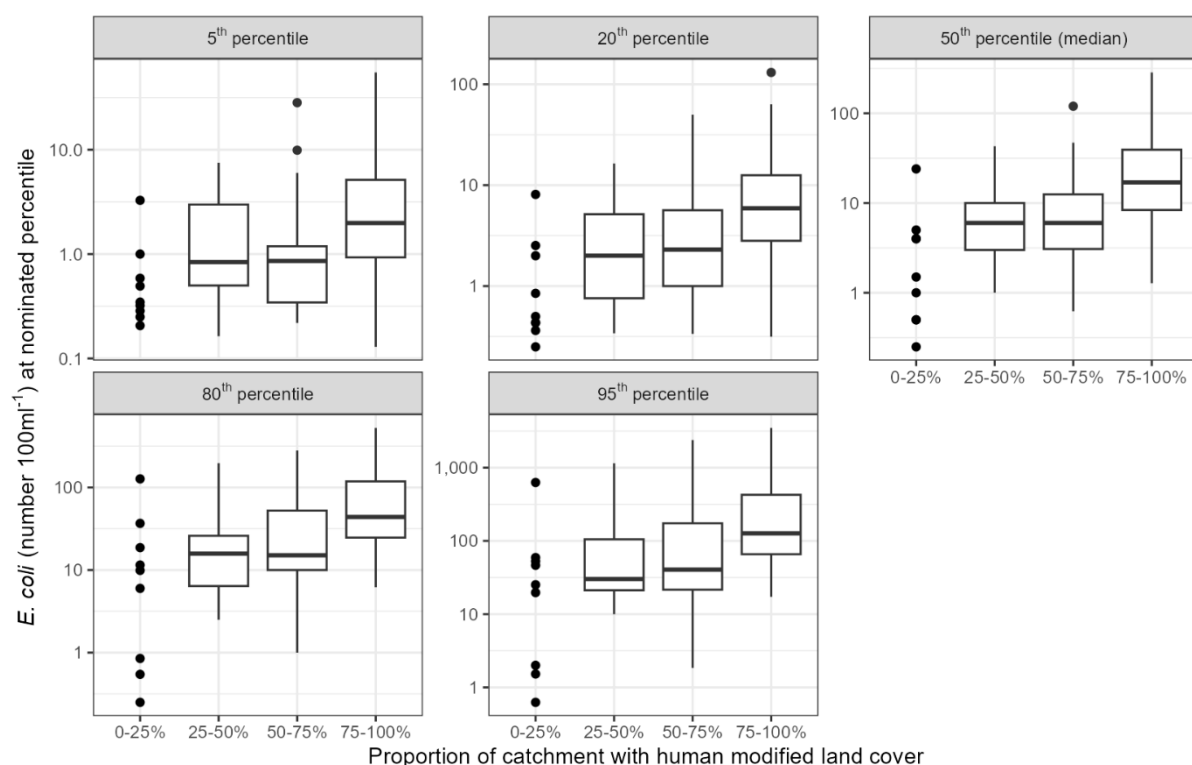
**Figure 4-4: Lake water quality state in lake geomorphology classes.** Geomorphological classes: glacial (G), dam (D), volcanic (V), aeolian (A), landslide (L), riverine (R), peat (P), and shoreline (S). Classes with  $\geq 10$  lakes are shown as box plots, and classes with  $< 10$  lakes are shown as points indicating individual lake median values. Box plots represent percentiles of lake medians. Horizontal lines are the medians of lake medians, boxes mark interquartile ranges, and whiskers indicate  $1.5 \times$  the interquartile ranges. Lake medians outside of the range of whiskers are shown as black circles. Note log-scale on Y-axes.



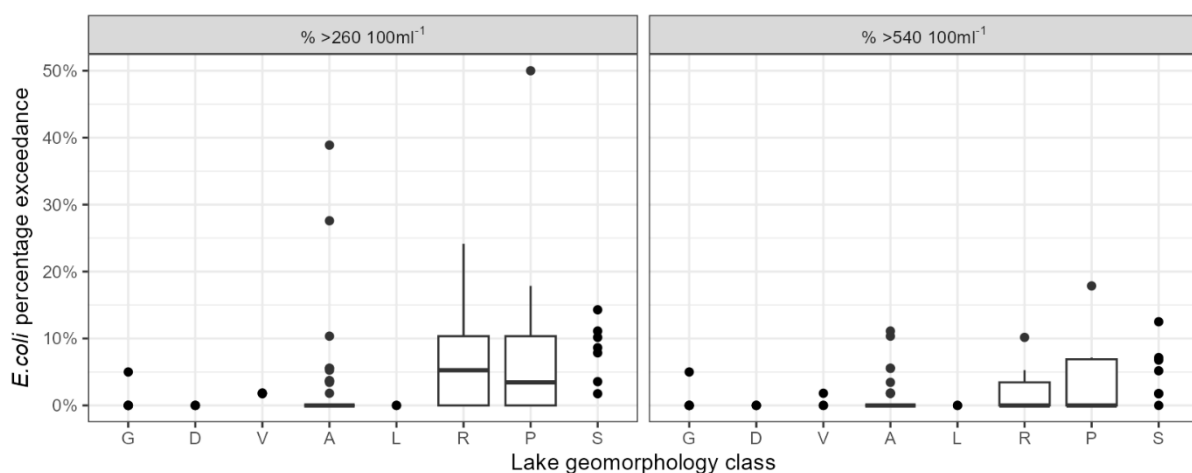
**Figure 4-5: Lake water quality state in human modified land cover classes.** Classes with  $\geq 10$  lakes are shown as box plots, and classes with  $< 10$  lakes are shown as points indicating individual lake median values. Box plots represent percentiles of lake medians. Horizontal lines are the medians of lake medians, boxes mark interquartile ranges, and whiskers indicate  $1.5 \times$  the interquartile ranges. Lake medians outside of the range of whiskers are shown as black circles. Note log-scale on Y-axes.



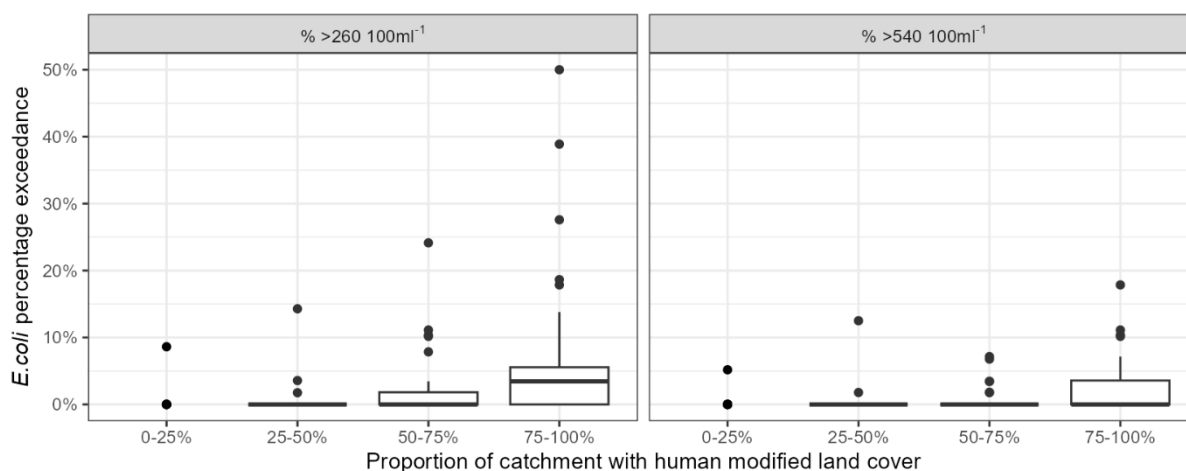
**Figure 4-6: *E. coli* concentrations in lake geomorphological classes.** Geomorphological classes: glacial (G), dam (D), volcanic (V), aeolian (A), landslide (L), riverine (R), peat (P), and shoreline (S). Classes with  $\geq 10$  lakes are shown as box plots, and classes with  $< 10$  lakes are shown as points indicating individual lake median values. Box plots represent percentiles of lake medians. Horizontal lines are the medians of lake medians, boxes mark interquartile ranges, and whiskers indicate  $1.5 \times$  the interquartile ranges. Lake medians outside of the range of whiskers are shown as black circles. Note log-scale on Y-axes.



**Figure 4-7: *E. coli* concentrations in human modified land cover classes.** Classes with  $\geq 10$  lakes are shown as box plots, and classes with  $< 10$  lakes are shown as points indicating individual lake median values. Box plots represent percentiles of lake medians. Horizontal lines are the medians of lake medians, boxes mark interquartile ranges, and whiskers indicate  $1.5\times$  the interquartile ranges. Lake medians outside of the range of whiskers are shown as black circles. Note log-scale on Y-axes.



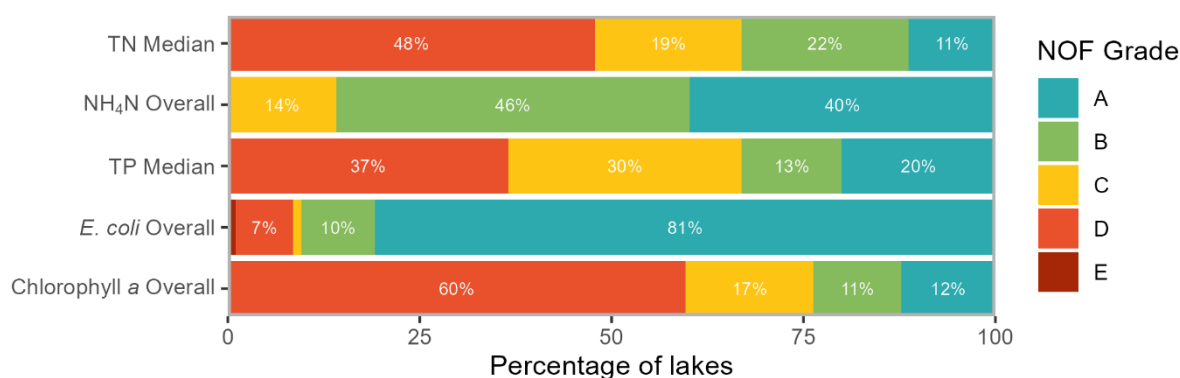
**Figure 4-8: *E. coli* percent exceedance in lake geomorphological classes.** Geomorphological classes: glacial (G), dam (D), volcanic (V), aeolian (A), landslide (L), riverine (R), peat (P), and shoreline (S). Classes with  $\geq 10$  lakes are shown as box plots, and classes with  $< 10$  lakes are shown as points indicating individual lake percentage exceedance values. Box plots represent percentiles of lake percentage exceedances. Horizontal lines are the medians of lake percentage exceedances, boxes mark interquartile ranges, and whiskers indicate  $1.5\times$  the interquartile ranges. Lake percentage exceedances outside of the range of whiskers are shown as black circles.



**Figure 4-9: *E. coli* percent exceedance in human modified land cover classes.** Classes with  $\geq 10$  lakes are shown as box plots, and classes with  $< 10$  lakes are shown as points indicating individual lake percentage exceedance values. Box plots represent percentiles of lake percentage exceedances. Horizontal lines are the medians of lake percentage exceedances, boxes mark interquartile ranges, and whiskers indicate  $1.5 \times$  the interquartile ranges. Lake percentage exceedances outside of the range of whiskers are shown as black circles.

## 4.2 NOF grades

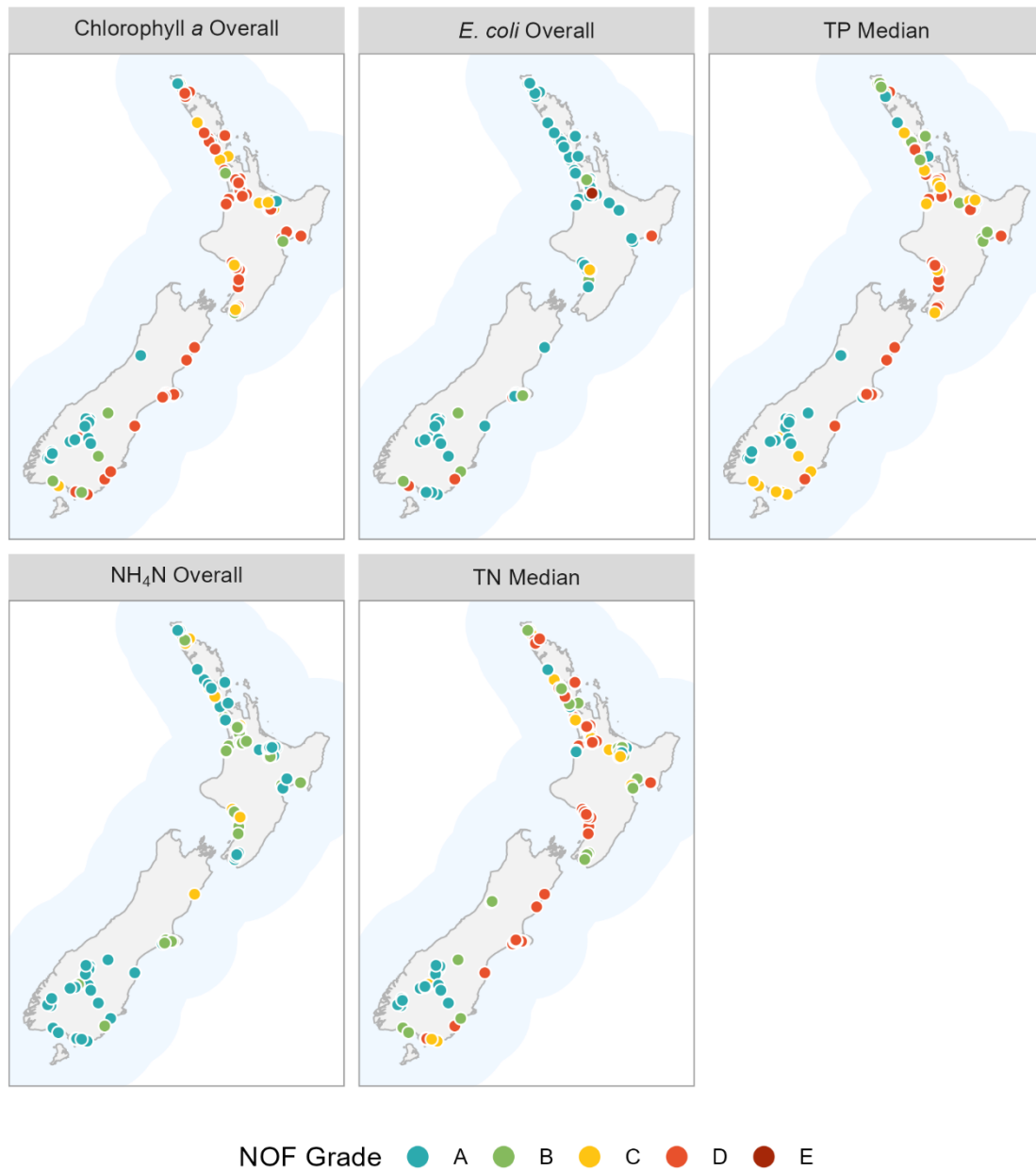
Figure 4-10 and Table 4-3 provide summaries of water quality grades for each NPS-FM attribute by demonstrating the number and percentage of lakes that are classified in each NOF grade. Figure 4-11 and Figure 4-12 provide maps for each attribute showing the sites coloured by their evaluated state grade. The percentage of sites belonging to each grade, by variable and geomorphology class are shown in Figure 4-13 and Figure 4-15, and by upstream human modified land cover class in Figure 4-14 and Figure 4-16. For the *E. coli* overall numeric attribute, 8% of lakes were graded D or E. These exceedances all occurred in shoreline, riverine, peat and aeolian lakes. No lakes were graded D or E for  $\text{NH}_4\text{-N}$ . Several lakes were below the bottom line for the phytoplankton (trophic state) attribute, with 60% of sites assigned a D grade for overall chlorophyll *a* attribute. For the TN and TP attributes, 49% and 37% of sites were graded D, respectively. D grades occurred for TP and TN in all geomorphological classes, except for glacial, and across all upstream human modified land cover classes. However, a higher proportion of lakes in catchments with extensive human modified land cover were graded D for both TP and TN. There were also greater proportions of lakes with D grades for the *E. coli* overall numeric attribute and the overall chlorophyll *a* attribute with greater proportion of catchment with human modified cover.



**Figure 4-10: Stacked bar chart showing the percentage of lakes assigned each NOF grade overall NOF attribute.**

**Table 4-3: Summary of the number of lakes and sites (in brackets, where this is different from the number of lakes) in each NOF grade.** Cells shown in grey indicate grades that are below the NOF national bottom line.

Numeric attribute state	NOF Grade				
	A	B	C	D	E
Chlorophyll a Overall	14 (25)	16	20 (25)	68 (82)	
Chlorophyll a Maximum	20 (31)	15 (17)	19 (24)	63 (76)	
Chlorophyll a Median	16 (27)	23 (24)	28 (36)	50 (61)	
<i>E. coli</i> Overall	81 (97)	10 (11)	2	7 (8)	1
<i>E. coli</i> % >260	90 (113)	3	1	2	0
<i>E. coli</i> % >540	82 (98)	12 (14)	6 (7)	0	0
<i>E. coli</i> Median	92 (116)	0	0	2	1
<i>E. coli</i> 95 <sup>th</sup> percentile	84 (100)	11 (12)	1	5 (6)	0
NH <sub>4</sub> -N Overall	48 (68)	52 (60)	16	0	
NH <sub>4</sub> -N Median	97 (128)	15	1	0	
NH <sub>4</sub> -N 95 <sup>th</sup> percentile	48 (68)	52 (60)	16	0	
TN Median	13 (24)	26 (34)	22 (24)	55 (65)	
TP Median	23 (35)	17 (18)	35 (43)	42 (51)	



**Figure 4-11: Maps showing lake site NOF overall attribute grades.**

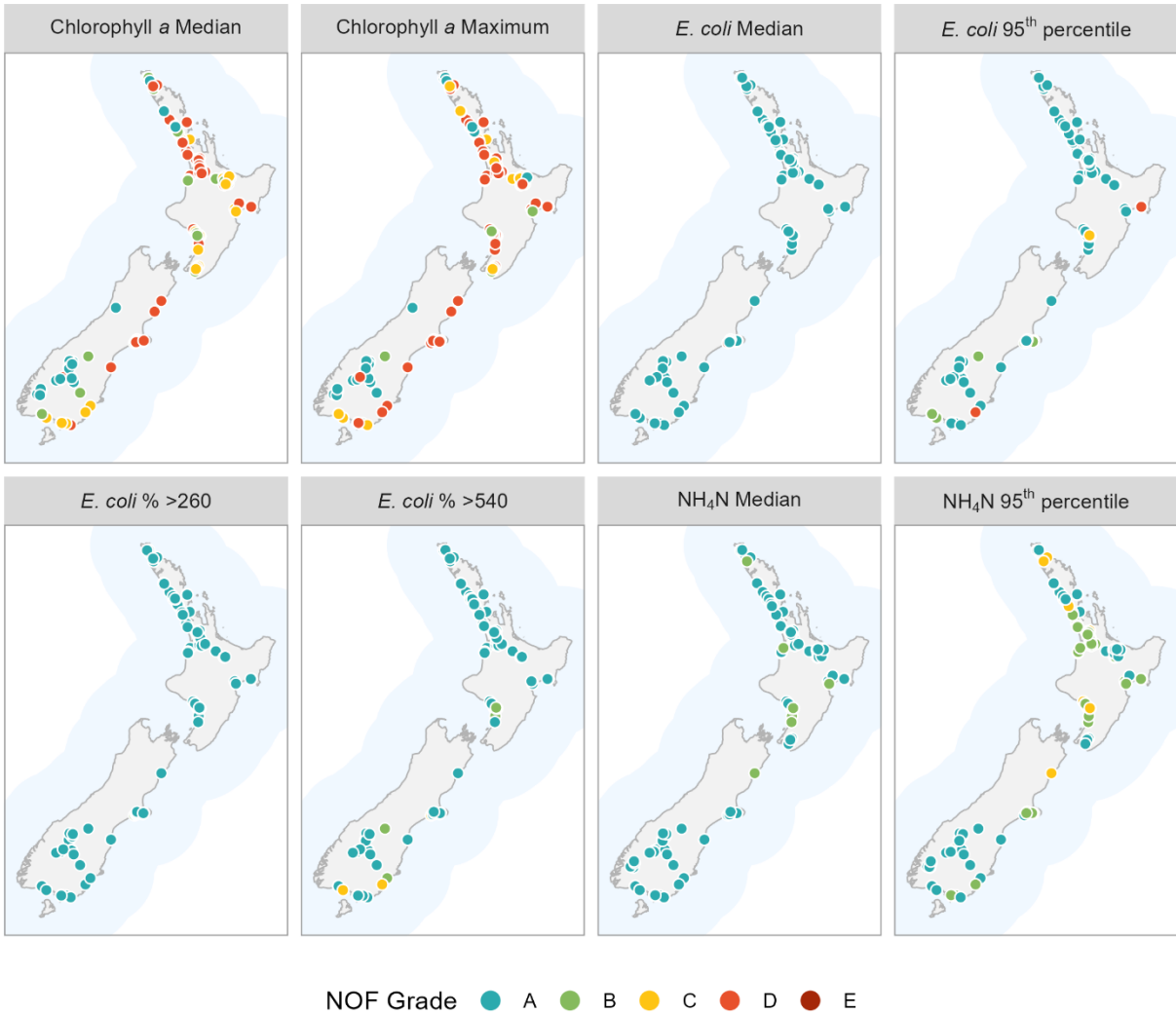
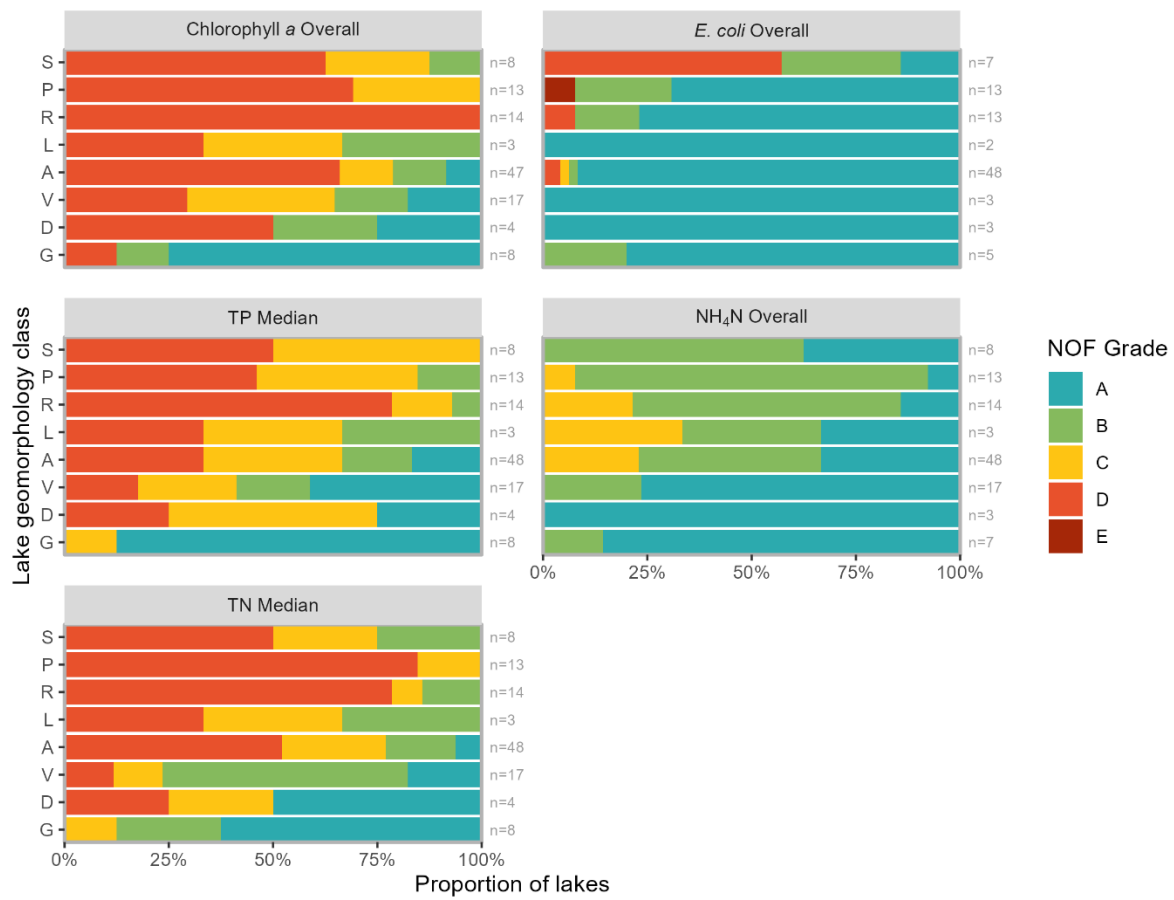
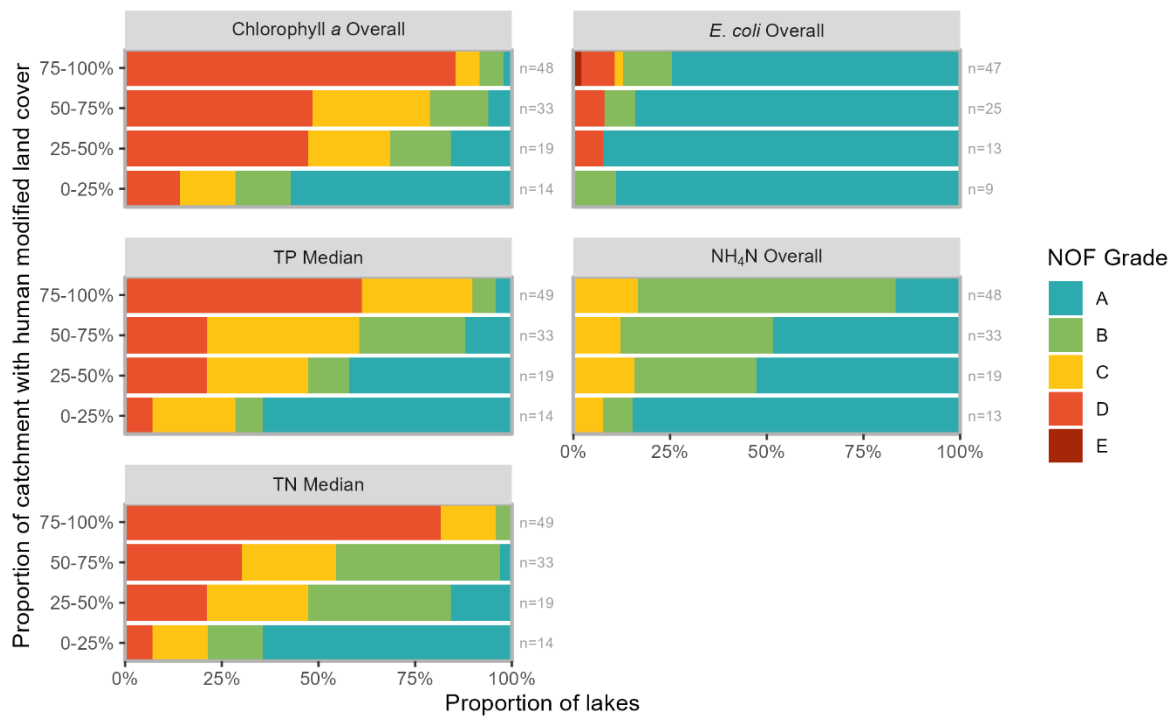


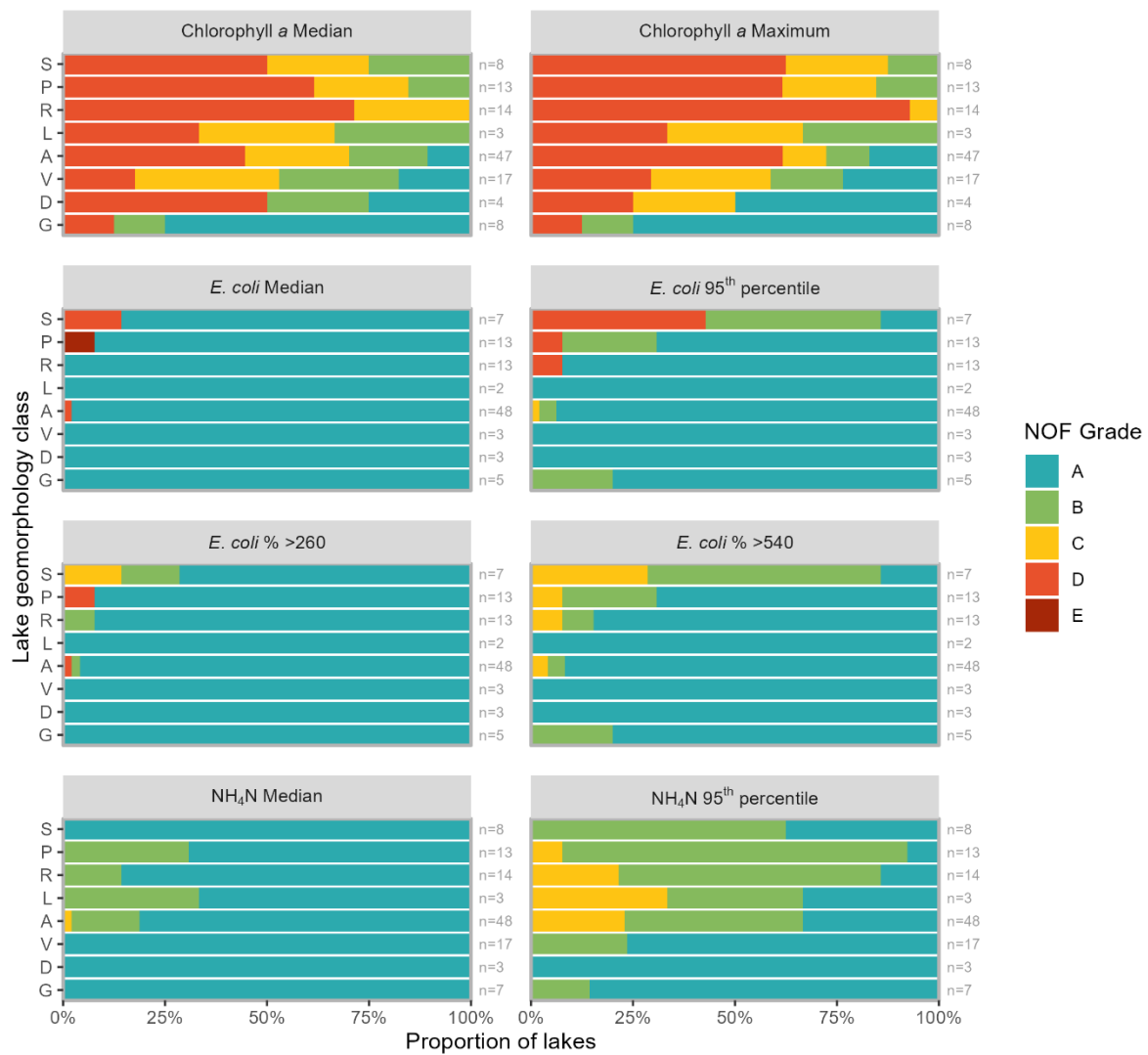
Figure 4-12: Maps showing lake site grades for NOF attribute components.



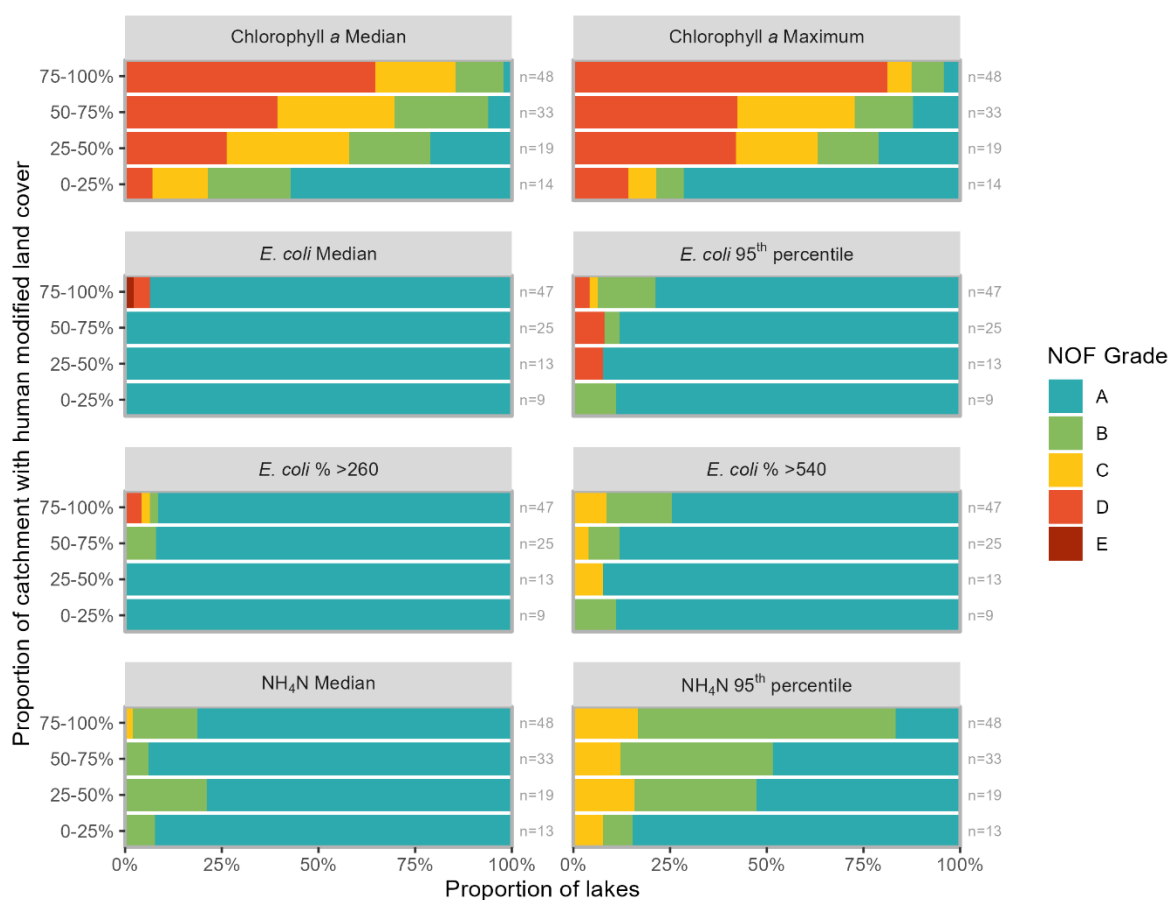
**Figure 4-13: Stacked bar charts showing the percentage of lakes assigned to each NOF grade), by lake geomorphological class for NOF overall attributes.** n indicates the number lakes in each class. Geomorphological classes: glacial (G), dam (D), volcanic (V), aeolian (A), landslide (L), riverine (R), peat (P), and shoreline (S).



**Figure 4-14: Stacked bar charts showing the percentage of lakes assigned to each NOF grade (median over sites within a lake), by lake upstream human modified land cover classes for NOF overall attributes. n indicates the number lakes in each class.**



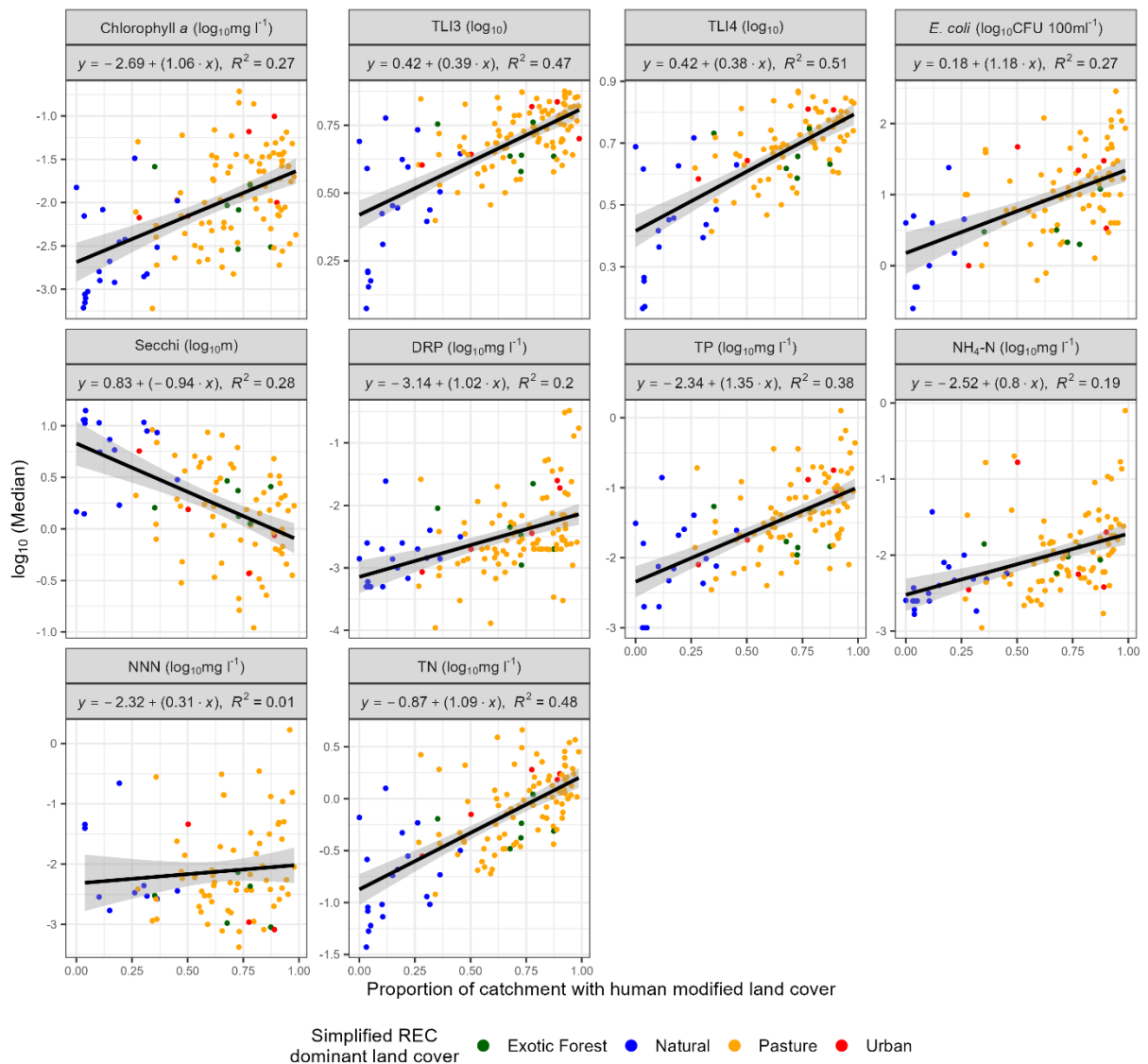
**Figure 4-15: Stacked bar charts showing the percentage of lakes assigned to each NOF grade (median over sites within a lake), by lake geomorphological class for NOF attribute components. n indicates the number lakes in each class. Geomorphological classes: glacial (G), dam (D), volcanic (V), aeolian (A), landslide (L), riverine (R), peat (P), and shoreline (S).**



**Figure 4-16: Stacked bar charts showing the percentage of lakes assigned to each NOF grade (median over sites within a lake), by lake upstream human modified land cover classes for NOF attribute components. n indicates the number lakes in each class.**

### 4.3 Relationships between water quality state and catchment land cover

The regression results indicated that the concentrations of each nutrient (except NNN) and *E. coli* increased, and Secchi decreased, with increasing proportions of upstream human modified land cover (Figure 4-17). Human modified land cover explained only 1% of variation in log-transformed NNN, but for the other variables human modified land cover explained 19%–51% of the variation in log-transformed water quality variables; these relationships were strongest for median TN, TP, TLI3 and TLI4. Colouring of the points in Figure 4-17 indicated that very few sampled lakes had dominant urban or exotic forest land cover. There was some evidence for sites with dominant urban land cover to be associated with high TLI and sites with dominant exotic forest land cover to be associated with lower MCI scores with respect to the fitted regression line. There was also a tendency for sites with dominant exotic forest land cover to be associated with low *E. coli* concentrations.



**Figure 4-17: Relationships between lake median water quality state and proportion of upstream catchment with human modified land cover.** Lakes are coloured based on their simplified REC dominant land cover class. Solid lines indicate least squares linear regression models. The shaded grey bands around the regression lines represent 95% confidence intervals for the fitted linear models.

## 5 Results – lake trends

### 5.1 Ten-year trends (2015–2024)

Between 30 and 84 lakes (and 45 and 108 sites) met the filtering rules for the 10-year trend analysis of water quality variables. The number of lakes varied widely by lake geomorphology class (Table 5-1) and included more lakes with greater upstream human modified land cover than not (Table 5-2). The qualifying lakes are sparsely and unevenly distributed on both the North and South Islands (Figure 5-1).

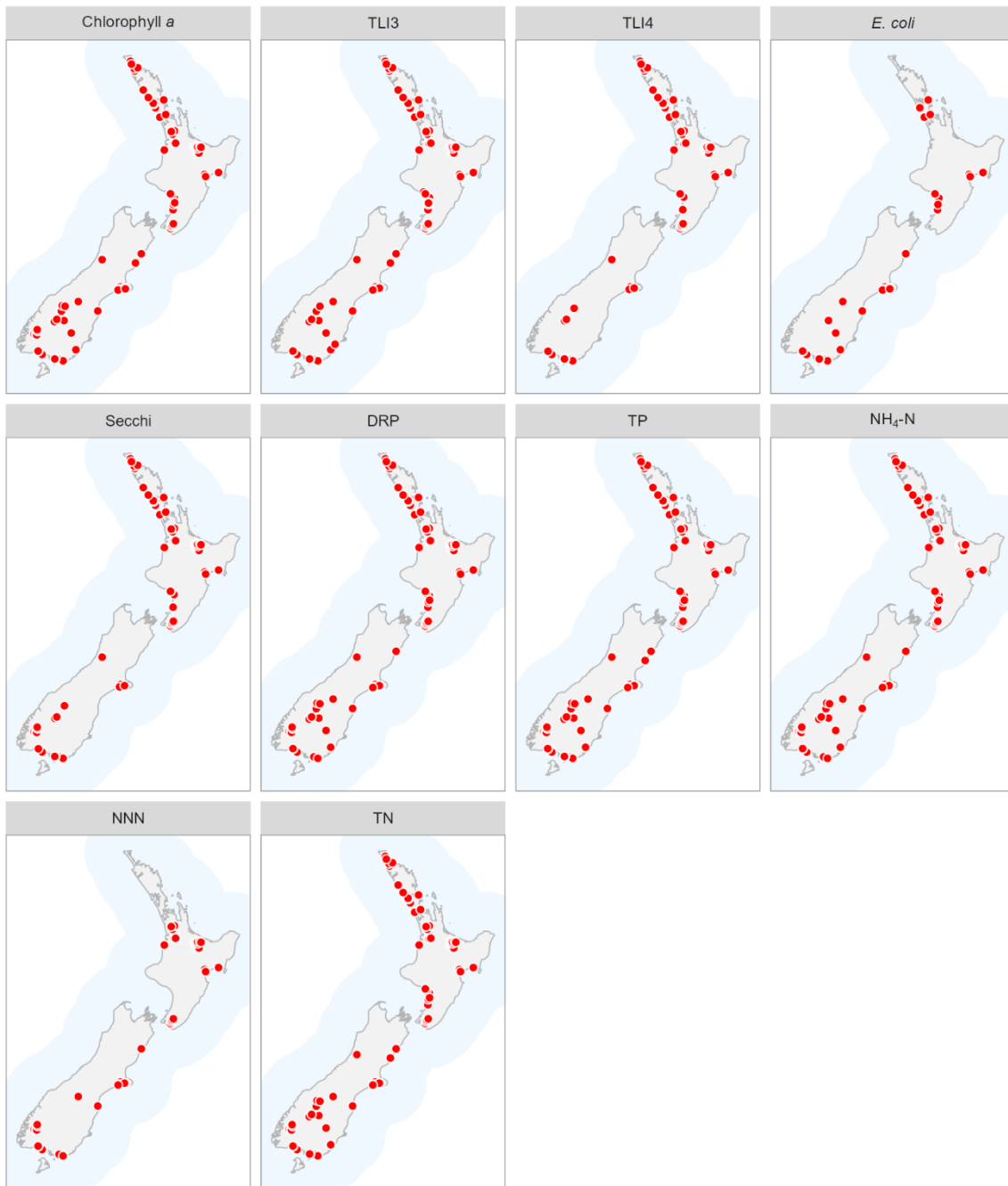
For the 10-year trend period, a large proportion of observations were censored at a small number of sites for some variables (chlorophyll *a*, Secchi, TN, TP), while others (DRP, NH<sub>4</sub>-N, *E. coli*, NNN) exhibited a larger proportion of censored observations at more sites compared to the other variables (Figure 5-2).

**Table 5-1: Number of lakes and sites (in brackets where different from the number of lakes) in the lake geomorphological classes included in the 10-year trend analyses.** The lake and site numbers shown refer to lakes or sites that met the site inclusion requirements in Section 3.2.1 (measurements were available for at least 80% of the years and at least 80% of seasons).

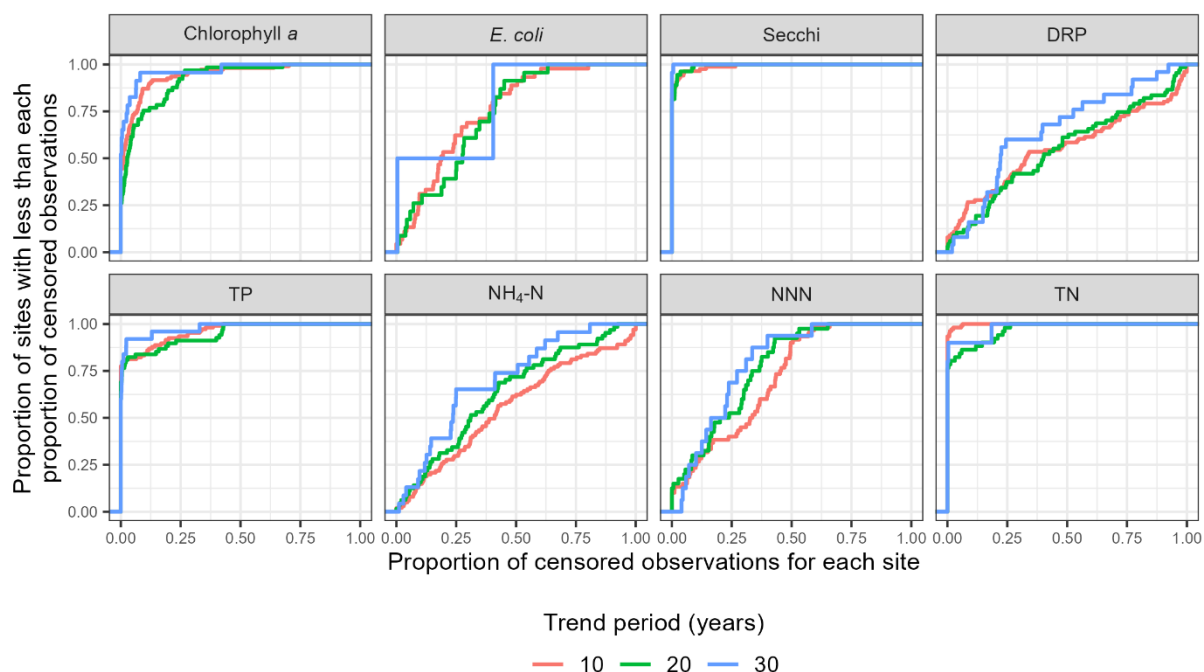
Variable	Number of lakes (and sites)								
	Total	Glacial	Dam	Volcanic	Aeolian	Landslide	Riverine	Peat	Shoreline
Chlorophyll <i>a</i>	80 (108)	8 (16)	4 (5)	12 (15)	30 (31)	2	9 (11)	7	8 (21)
TLI3	84 (106)	6 (10)	4 (5)	12 (15)	36 (37)	2	10 (12)	6	8 (19)
TLI4	61 (78)	4	1 (2)	12 (15)	23 (24)	2	7 (9)	6	6 (16)
<i>E. coli</i>	30 (45)	1	3 (4)	1	13 (14)	2	3	0	7 (20)
Secchi	64 (84)	6 (10)	1 (2)	12 (15)	23 (24)	2	7 (9)	7	6 (15)
DRP	64 (79)	2	3 (4)	12 (15)	28 (29)	2	6 (8)	4	7 (15)
TP	80 (107)	8 (17)	4 (5)	12 (15)	30 (31)	2	9 (11)	7	8 (19)
NH <sub>4</sub> -N	73 (88)	3	3 (4)	12 (15)	30 (31)	2	9 (11)	7	7 (15)
NNN	42 (60)	3 (6)	1 (2)	11 (14)	3 (4)	2	8 (10)	7	7 (15)
TN	79 (106)	8 (17)	4 (5)	12 (15)	30 (31)	2	9 (11)	6	8 (19)

**Table 5-2: Number of lakes and sites (in brackets where different from the number of lakes) in human modified land cover classes included in the 10-year trend analyses.** The lake and site numbers shown refer to lakes or sites that met the site inclusion requirements in Section 3.2.1 (measurements were available for at least 80% of the years and at least 80% of seasons).

Variable	Number of lakes (and sites)				
	Total	0–25%	25–50%	50–75%	75–100%
Chlorophyll <i>a</i>	80 (108)	13 (23)	14 (17)	23 (34)	30 (34)
TLI3	84 (106)	11 (17)	14 (17)	24 (33)	35 (39)
TLI4	61 (78)	7 (8)	13 (16)	22 (31)	19 (23)
<i>E. coli</i>	30 (45)	5 (7)	5 (6)	8 (16)	12 (16)
Secchi	64 (84)	9 (13)	13 (16)	22 (31)	20 (24)
DRP	64 (79)	8 (10)	14 (17)	19 (25)	23 (27)
TP	80 (107)	13 (24)	14 (17)	23 (32)	30 (34)
NH <sub>4</sub> -N	73 (88)	8 (10)	14 (17)	22 (28)	29 (33)
NNN	42 (60)	5 (10)	9 (12)	13 (19)	15 (19)
TN	79 (106)	13 (24)	14 (17)	23 (32)	29 (33)



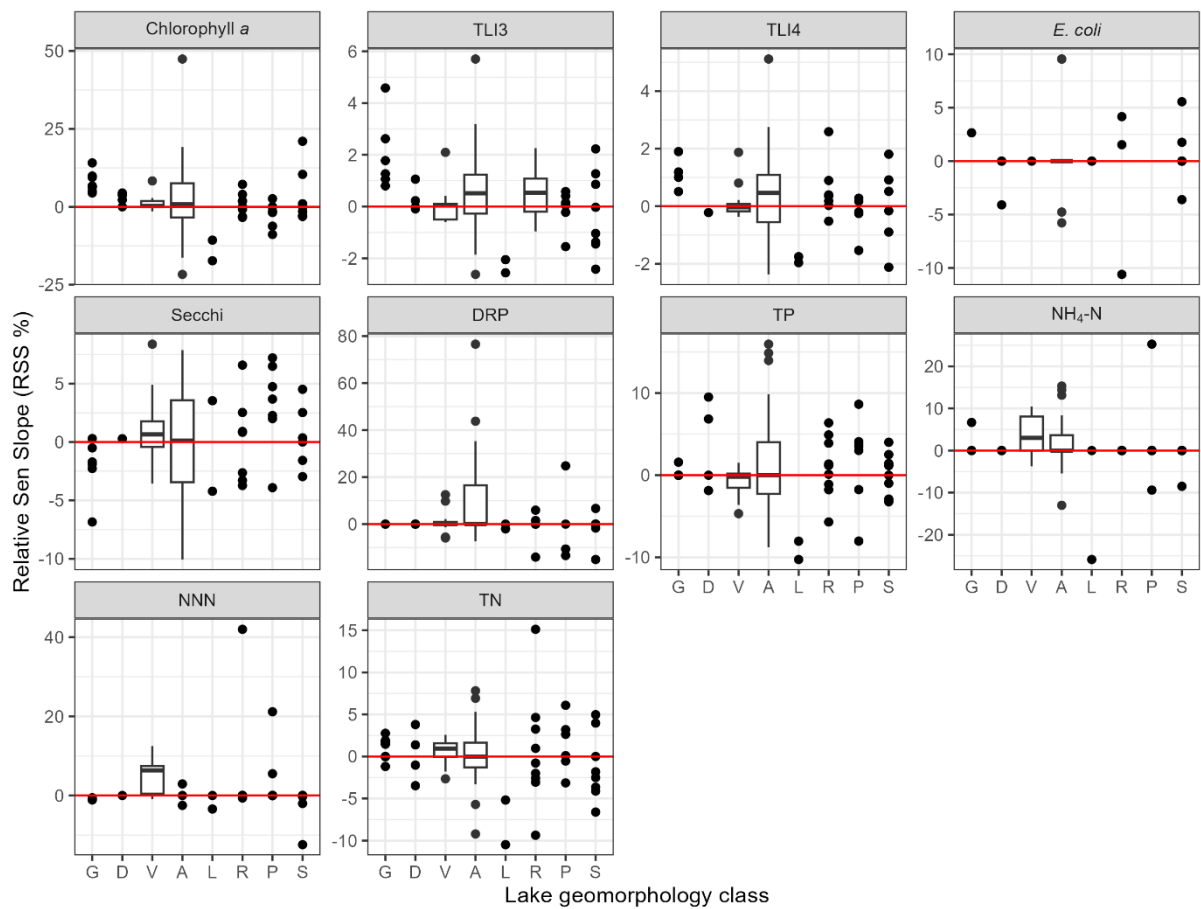
**Figure 5-1: Locations of lake sites included in the 10-year trend analyses by water quality variable.**



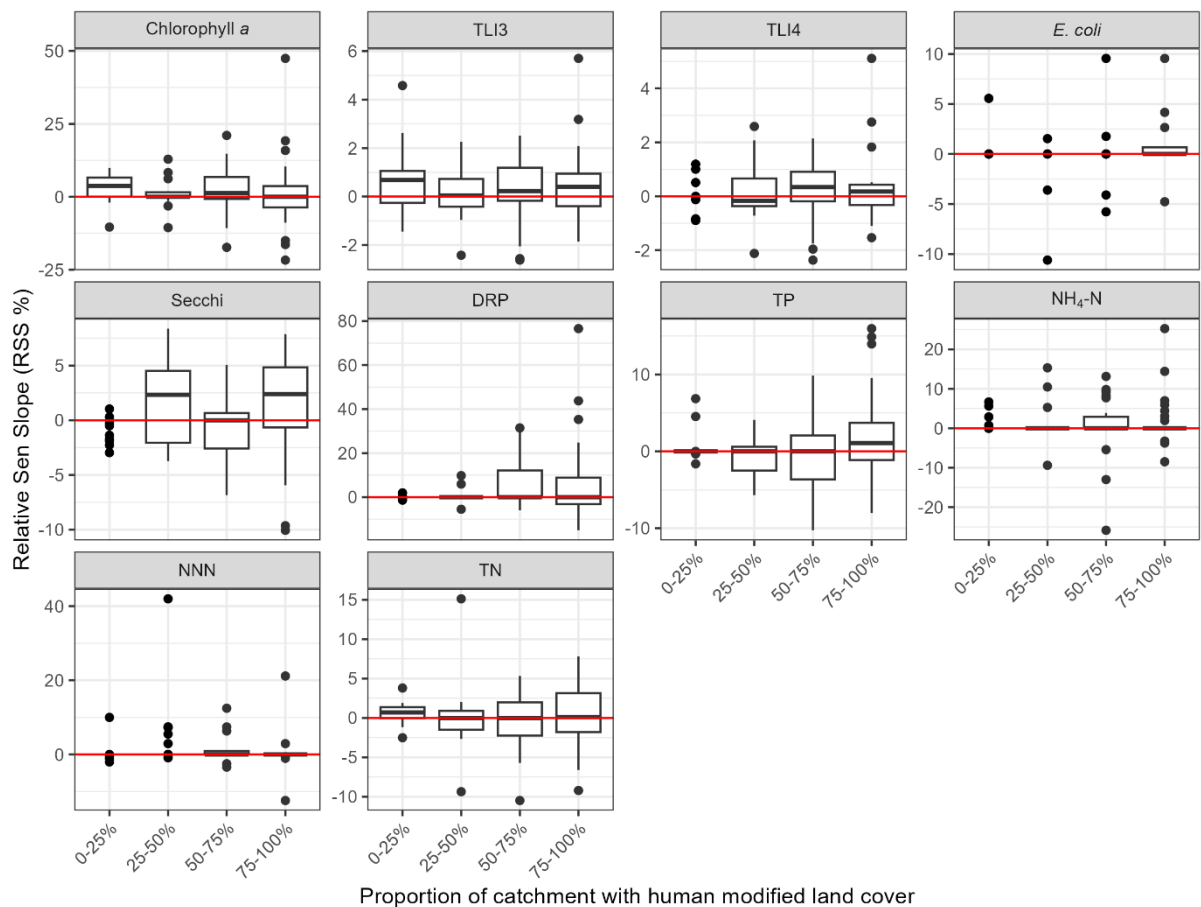
**Figure 5-2: Proportion of censored data for each site × variable combination for which trend was assessed.** Note TLI3 and TLI4 are not included as the aggregate annual values are not censored.

### 5.1.1 Trend rate

Box-and-whisker plots summarise the assessed trend rates for each lake and water quality variable for the 10-year period from 2015–2024 in each lake geomorphological class (Figure 5-3) and each upstream human modified land cover class (Figure 5-4). All estimated lake trend rates are included in these plots, irrespective of the level of confidence in the assessment (as defined in Section 3.2.3). The plots indicate that geomorphological and land cover classes did not account for a substantial amount of the variation in trends between lakes for any variable; this is in contrast with the state analyses of water quality variables, where water quality clearly varied between some geomorphological and land cover classes (Figure 4-4, Figure 4-6, Figure 4-8). Further interpretation of differences in trend magnitudes between each lake geomorphological class was limited because there were only two classes that contained more than 10 lakes.



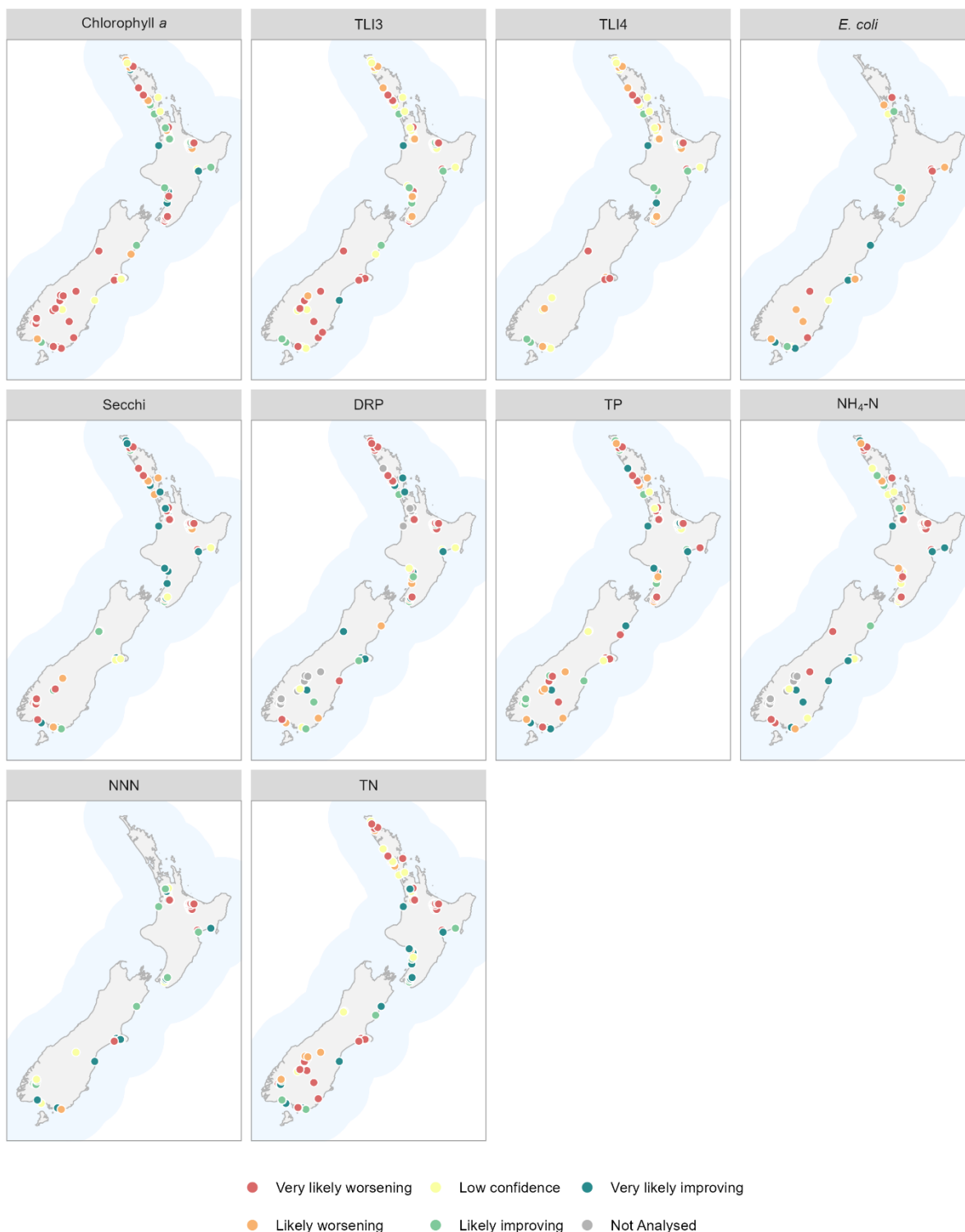
**Figure 5-3: Summary of 10-year lake trend rates for the water quality variables, in lake geomorphological classes.** Geomorphological classes: glacial (G), dam (D), volcanic (V), aeolian (A), landslide (L), riverine (R), peat (P), and shoreline (S). Classes with  $\geq 10$  lakes are shown as box plots, and classes with  $< 10$  lakes are shown as points indicating individual lake values. Box plots represent percentiles of lake relative Sen slopes. Horizontal lines are the medians of lake relative Sen slopes, boxes mark interquartile ranges, and whiskers indicate  $1.5 \times$  the interquartile ranges. Lake relative Sen slopes outside of the range of whiskers are shown as black circles.



**Figure 5-4: Summary of 10-year lake trend rates for the water quality variables in human modified land cover classes.** Classes with  $\geq 10$  lakes are shown as box plots, and classes with  $< 10$  lakes are shown as points indicating individual lake values. Box plots represent percentiles of lake relative Sen slopes. Horizontal lines are the medians of lake relative Sen slopes, boxes mark interquartile ranges, and whiskers indicate  $1.5 \times$  the interquartile ranges. Lake relative Sen slopes outside of the range of whiskers are shown as black circles.

**5.1.2 Trend direction**

The levels of confidence listed in Table 3-2 were used to categorise the confidence that 10-year trends were improving for each site  $\times$  variable combination. The spatial distributions of categorised individual sites are shown in Figure 5-5.

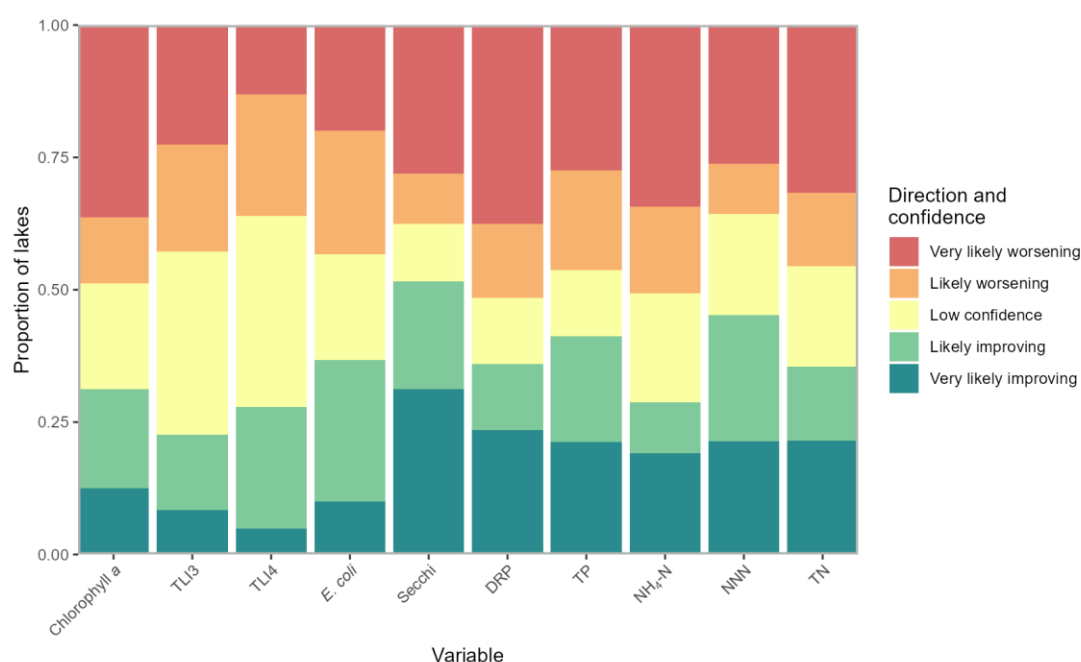


**Figure 5-5: Lake water quality monitoring sites categorised by the confidence that the 10-year trend was improving ( $C_i$ ) for each variable.**  $C_i$  is expressed using the confidence categories in Table 3-2. Only sites that met the sampling requirements outlined in Section 3.2.1 are shown in the figure.

### 5.1.3 Aggregate Trends

Figure 5-6 shows the proportions of lakes belonging to each of the five categorical levels of confidence for  $C_i$  defined in Table 3-2 for the 10-year trends. These plots provide a national-scale assessment of the relative proportions of lakes with improving versus worsening trends.

The national-scale proportions of increasing trends ( $P_i$ ) and their confidence intervals are summarised in Table 5-3. The 10-year  $P_i$  statistics ranged from 36–58%. Chlorophyll *a*, TLI3, DRP and  $NH_4$ -N had a majority (i.e.,  $P_i < 50\%$ ) of worsening trends at the 95% confidence level. For all other variables, the 95% confidence intervals for the  $P_i$  included 50% and we could not infer generality of increases or decreases across the analysed data for these variables.



**Figure 5-6: Summary plot representing the proportion of lakes with improving 10-year trends at each categorical level of confidence.** The plot shows the proportion of lakes with improving trends at levels of confidence defined in Table 3-2.

**Table 5-3: Proportions of improving 10-year lake trends ( $P_i$ ) by variable.**

Variable	Number of lakes	$P_i$ (%)	95% confidence interval for $P_i$ (%)
Chlorophyll <i>a</i>	80	40	31.8 – 48.2
TLI3	84	36.9	29.5 – 44.3
TLI4	61	45.1	31.5 – 58.7
<i>E. coli</i>	30	46.7	29.7 – 63.6
Secchi	64	57.8	48.6– 67.0
DRP	64	41.4	33.2 – 49.6
TP	80	48.8	39.0 – 58.5
$NH_4$ -N	73	36.3	22.9 – 49.7
NNN	42	54.8	25.1 – 84.4
TN	79	46.8	38.4 – 55.3

## 5.2 Twenty-year trends (2005–2024)

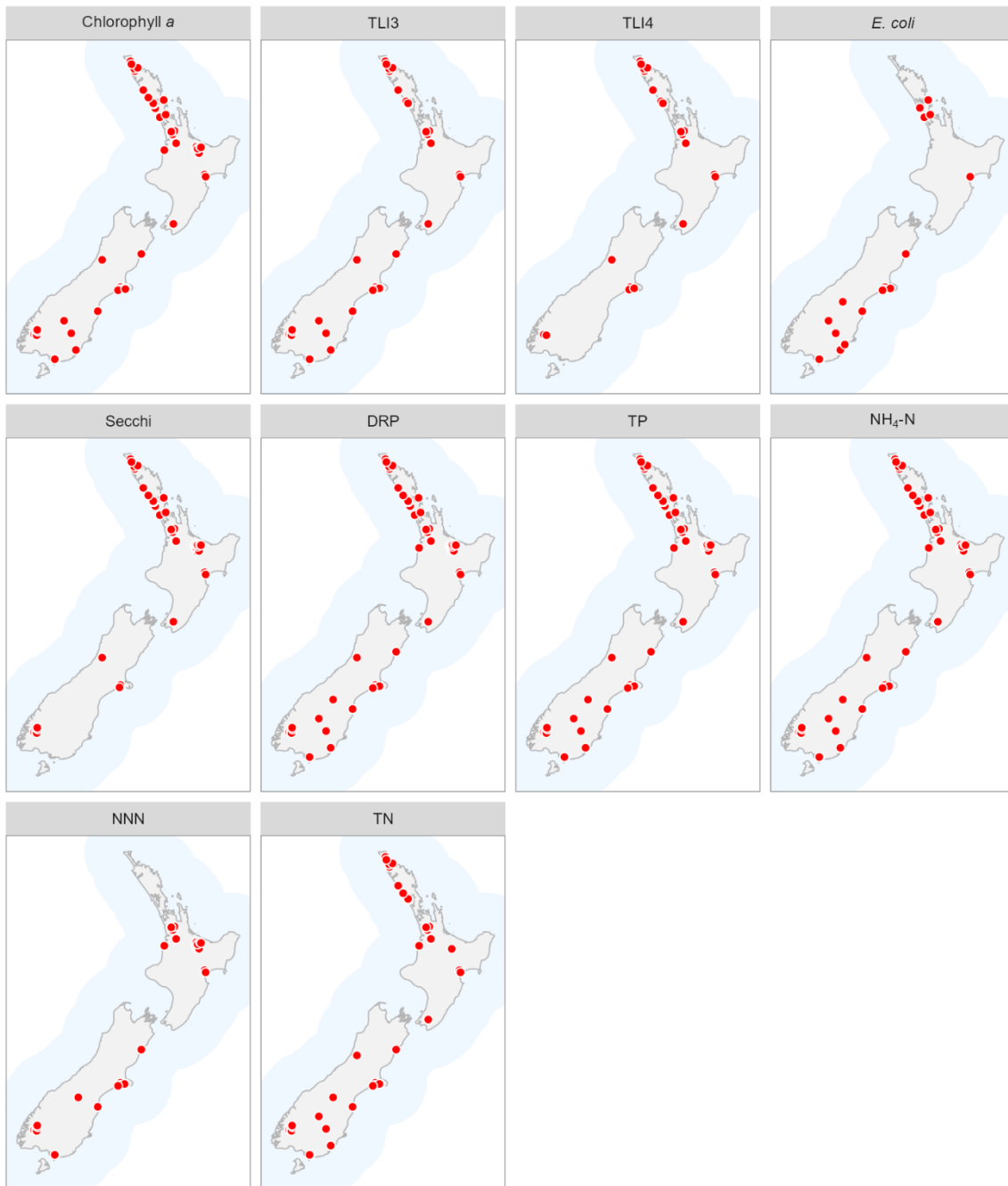
Between 17 and 54 lakes (and 23 and 68 sites) met the inclusion rules for the 20-year trend analysis (Table 5-4). Site-coverage for 20-year trend analysis was reduced compared to 10-year trend analysis. The locations of the lake sites included in the 20-year trend analyses are shown in Figure 5-7. The proportion of observations that were censored for each variable was similar for the 20-year period compared to the 10-year period (Figure 5-2).

**Table 5-4: Number of lakes and sites (in brackets where different from the number of lakes) in the lake geomorphology classes included in the 20-year trend analyses.** The site numbers shown refer to sites that met the site inclusion requirements in Section 3.2.1 (measurements were available for at least 80% of the years and at least 80% of seasons).

Variable	Number of lakes (and sites)								
	Total	Glacial	Dam	Volcanic	Aeolian	Landslide	Riverine	Peat	Shoreline
Chlorophyll <i>a</i>	52 (65)	3 (6)	2	11 (12)	19	2	8	3	4 (13)
TLI3	34 (44)	3 (6)	2		13	1	8	3	4 (11)
TLI4	25 (30)	2 (3)			12	1	6	3	1 (5)
<i>E. coli</i>	17 (23)	1	2	1	4	2	3		4 (10)
Secchi	44 (54)	3 (7)		11 (14)	18	2	6	3	1 (4)
DRP	45 (55)	4	2	12 (15)	14	2	6	1	4 (11)
TP	54 (68)	4 (8)	2	12 (15)	19	2	8	3	4 (11)
NH <sub>4</sub> -N	53 (64)	4 (7)	2	11 (12)	19	2	8	3	4 (11)
NNN	30 (40)	3 (6)		10 (11)	2	2	6	3	4 (10)
TN	40 (51)	4 (8)	2	1	16	2	8	3	4 (11)

**Table 5-5: Number of lakes and sites (in brackets where different from the number of lakes) in human modified land cover classes included in the 20-year trend analyses.** The site numbers shown refer to sites that met the site inclusion requirements in Section 3.2.1 (measurements were available for at least 80% of the years and at least 80% of seasons).

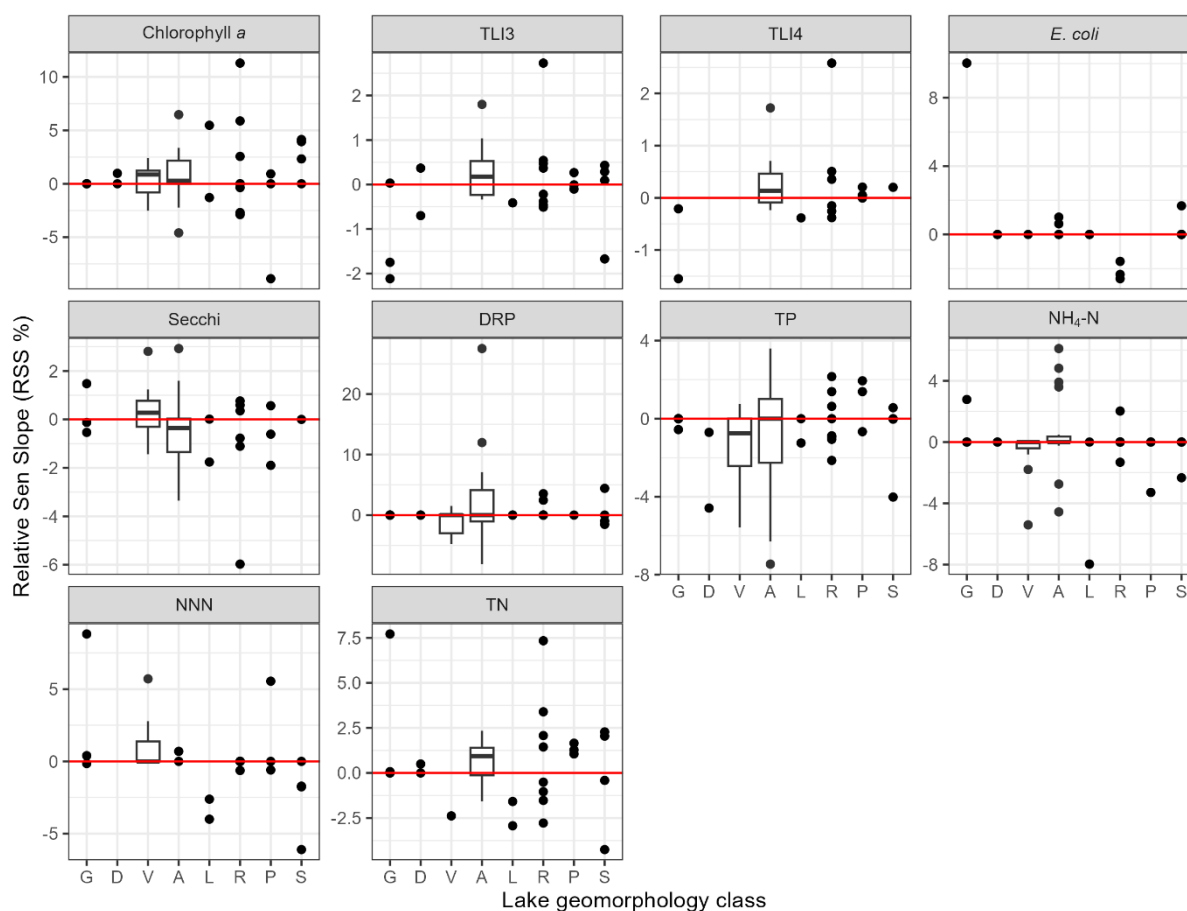
Variable	Number of lakes (and sites)				
	Total	0–25%	25–50%	50–75%	75–100%
Chlorophyll <i>a</i>	52 (65)	8 (11)	13	17 (23)	14 (18)
TLI3	34 (44)	5 (8)	7	10 (13)	12 (16)
TLI4	25 (30)	2 (3)	6	8	9 (13)
<i>E. coli</i>	17 (23)	3	3	5 (8)	6 (9)
Secchi	44 (54)	6 (10)	12	15 (18)	11 (14)
DRP	45 (55)	8	10	15 (21)	12 (16)
TP	54 (68)	8 (12)	13	18 (24)	15 (19)
NH <sub>4</sub> -N	53 (64)	8 (11)	13	17 (21)	15 (19)
NNN	30 (40)	4 (7)	7	9 (12)	10 (14)
TN	40 (51)	5 (9)	7	13 (16)	15 (19)



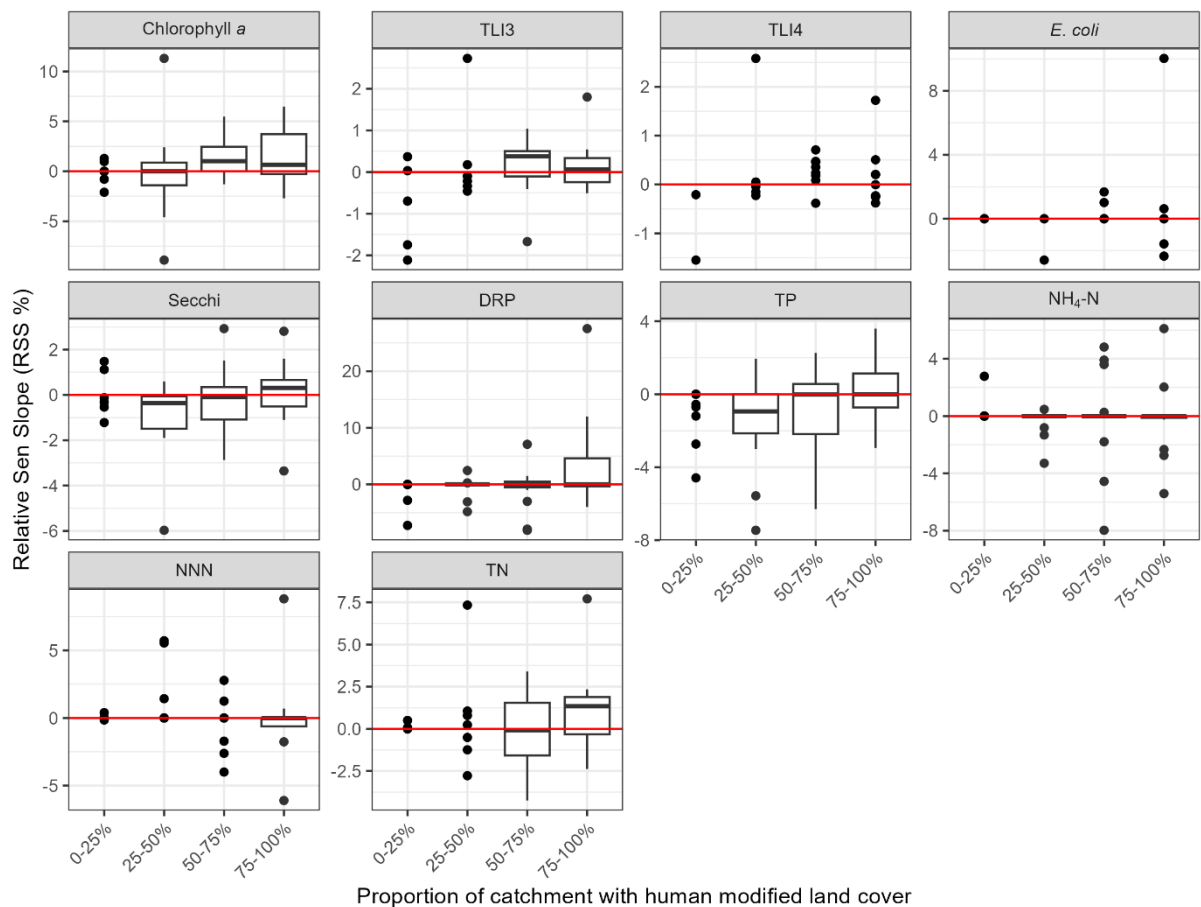
**Figure 5-7: Locations of lake sites included in the 20-year trend analyses by water quality variable.**

### 5.2.1 Trend rate

Box-and-whisker plots summarise the assessed trend rates for each lake water quality variable for the 20-year period from 2005–2024 in each lake geomorphological class (Figure 5-8) and each upstream human modified land cover class (Figure 5-9). All estimated trend rates are included in these plots, irrespective of the level of confidence in the assessment (see Section 3.2.3).



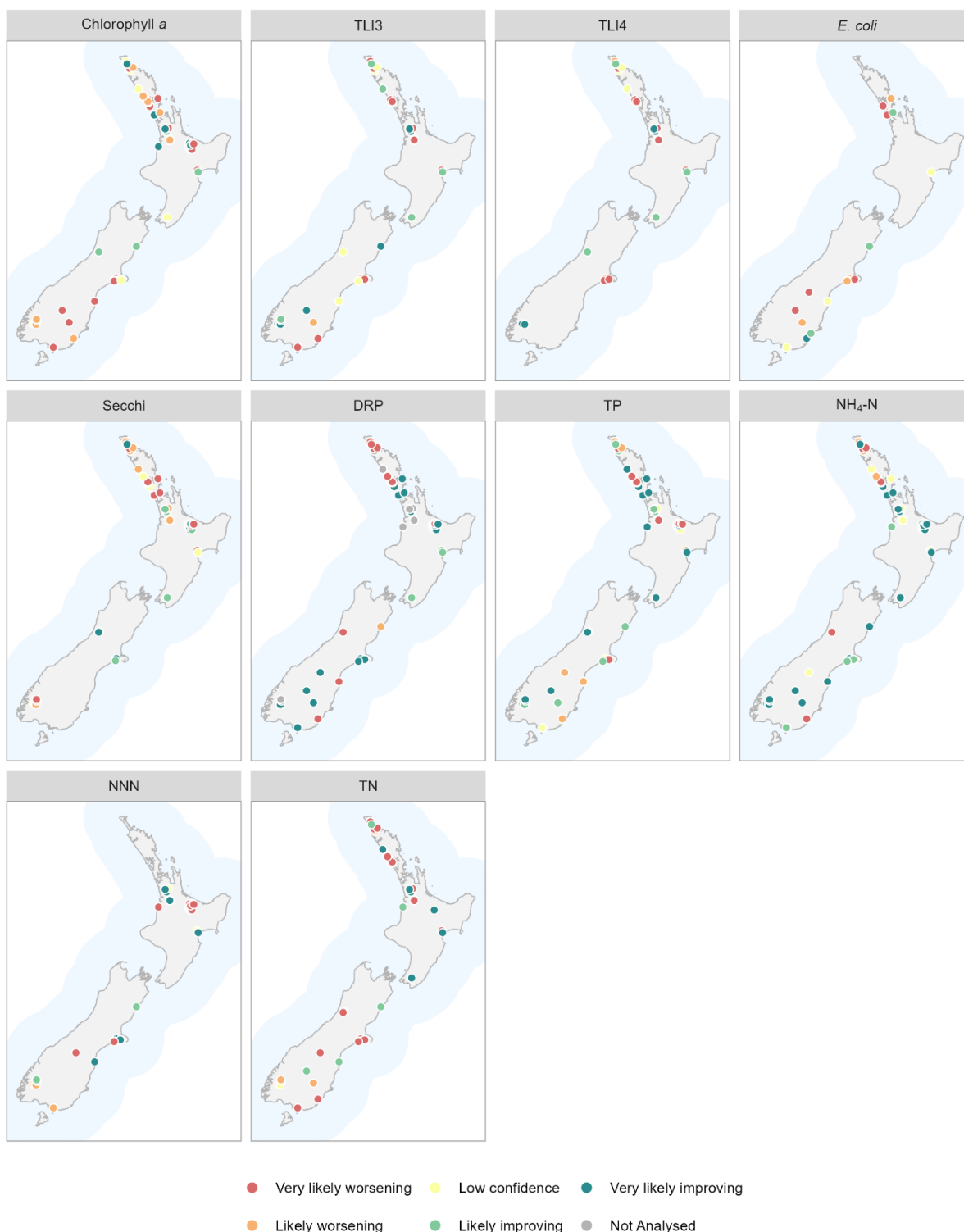
**Figure 5-8: Summary of 20-year lake trend rates for the water quality variables, in lake geomorphological classes.** Geomorphological classes: glacial (G), dam (D), volcanic (V), aeolian (A), landslide (L), riverine (R), peat (P), and shoreline (S). Classes with  $\geq 10$  lakes are shown as box plots, and classes with  $< 10$  lakes are shown as points indicating individual lake values. Box plots represent percentiles of lake relative Sen slopes. Horizontal lines are the medians of lake relative Sen slopes, boxes mark interquartile ranges, and whiskers indicate  $1.5 \times$  the interquartile ranges. Lake relative Sen slopes outside of the range of whiskers are shown as black circles.



**Figure 5-9: Summary of 20-year lake trend rates for the water quality variables in human modified land cover classes.** Classes with  $\geq 10$  lakes are shown as box plots, and classes with  $< 10$  lakes are shown as points indicating individual lake values. Box plots represent percentiles of lake relative Sen slopes. Horizontal lines are the medians of lake relative Sen slopes, boxes mark interquartile ranges, and whiskers indicate  $1.5 \times$  the interquartile ranges. Lake relative Sen slopes outside of the range of whiskers are shown as black circles.

**5.2.2 Trend direction**

The levels of confidence listed in Table 3-2 were used to categorise the confidence that 20-year trends were improving for each site  $\times$  variable combination. The spatial distributions of categorised individual sites are shown in Figure 5-10.

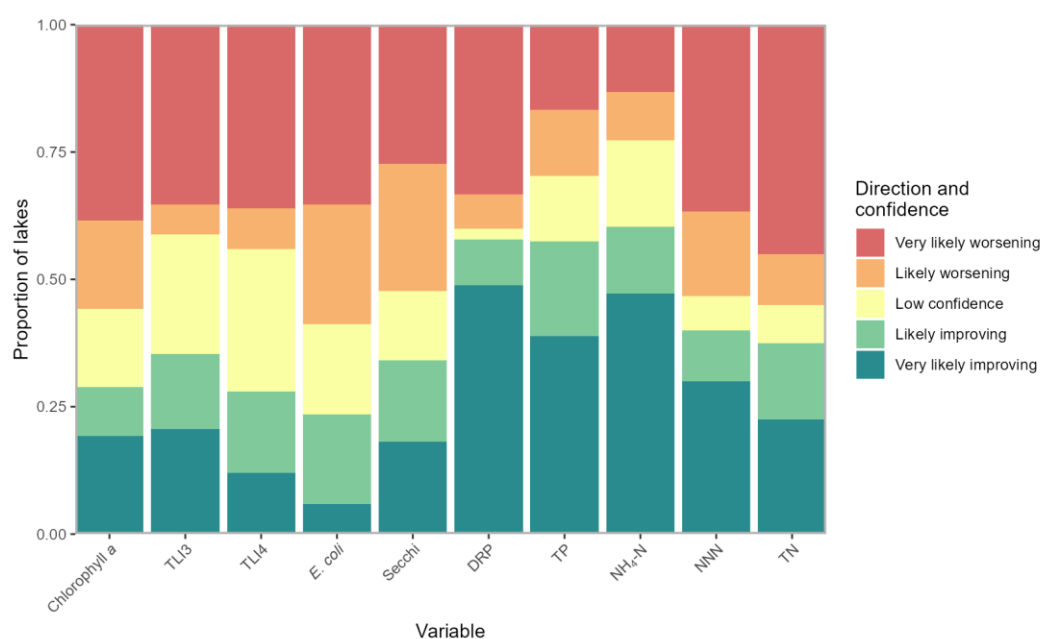


**Figure 5-10: Lake water quality monitoring sites categorised by the confidence that the 20-year trend was improving ( $C_i$ ) for each variable.**  $C_i$  is expressed using the confidence categories in Table 3-2. Only sites that met the sampling requirements outlined in Section 3.2.1 are shown in the figure.

### 5.2.3 Aggregate Trends

Figure 5-11 shows the proportions of lakes belonging to each of the five categorical levels of confidence for  $C_i$  defined in Table 3-2 for the 20-year, raw trends. These plots provide national-scale assessments of the relative proportions of lakes with improving versus worsening trends.

The national-scale proportions of improving trends ( $P_i$ ) and their confidence intervals are summarised in Table 5-6. The 10-year  $P_i$  statistics ranged from 24–69%. Chlorophyll *a*, *E. coli* and TN had a majority (i.e.,  $P_i < 50\%$ ) of worsening trends at the 95% confidence level. DRP, TP and  $NH_4-N$  as a majority (i.e.,  $P_i > 50\%$ ) of improving trends at the 95% confidence level. For all other variables (TLI3, TLI4, Secchi, NNN), the 95% confidence intervals included 50% and we could not infer generality of increases or decreases across the analysed data for these variables.



**Figure 5-11: Summary plot representing the proportion of lakes with improving 20-year time-period trends at each categorical level of confidence.** The plot shows the proportion of lakes with improving trends at levels of confidence defined in Table 3-2.

**Table 5-6: Proportions of improving trends ( $P_i$ ) for 20-year time period.**

Variable	Number of lakes	$P_i$ (%)	95% confidence interval for $P_i$ (%)
Chlorophyll <i>a</i>	52	36.5	26.7 – 46.4
TLI3	34	45.6	29.9 – 61.2
TLI4	25	42	25.5 – 58.5
<i>E. coli</i>	17	23.5	4.3 – 42.8
Secchi	44	43.2	26.9 – 59.5
DRP	45	57.8	51.0 – 64.6
TP	54	68.5	58.5 – 78.5
$NH_4-N$	53	67.9	56.6 – 79.2
NNN	30	43.3	24.6 – 62.0
TN	40	40	32.8 – 47.2

### 5.3 Thirty-year trends (1995–2024)

Between two and 15 lakes met the inclusion rules for the 30-year trend analysis for each of the water quality variables. Site-coverage for 30-year trend analysis was considerably reduced compared to 10-year trend analysis. The number of lakes varied widely by lake class (Table 5-7). The qualifying lakes were limited to three regions: Bay of Plenty (maximum of nine lakes), Waikato (two lakes), Otago (maximum five lakes) and Auckland (three lakes) (Figure 5-12). For the 30-year period, TLI was only available in Bay of Plenty. The proportion of observations that were censored for each variable was similar for the 30-year period compared to the 20-year period (Figure 5-2).

**Table 5-7: Number of lakes and sites (in brackets where different from the number of lakes) in the lake geomorphology classes included in the 30-year trend analyses.** The site numbers shown refer to sites that met the site inclusion requirements in Section 3.2.1 (measurements were available for at least 80% of the years and at least 80% of seasons).

Variable	Number of lakes (and sites)								
	Total	Glacial	Dam	Volcanic	Aeolian	Landslide	Riverine	Peat	Shoreline
Chlorophyll <i>a</i>	20 (23)		1	10	4		2	1	2 (5)
TLI3	7 (10)		1				3	1	2 (5)
TLI4	5 (8)						3	1	1 (4)
<i>E. coli</i>	2		1				1		
Secchi	19 (23)			11 (12)	4		2	1	1 (4)
DRP	20 (24)		1	10 (11)	4		3		2 (5)
TP	21 (25)		1	10 (11)	4		3	1	2 (5)
NH <sub>4</sub> -N	20 (23)		1	9	4		3	1	2 (5)
NNN	13 (16)			8			2	1	2 (5)
TN	7 (10)		1				3	1	2 (5)

**Table 5-8: Number of lakes and sites (in brackets where different from the number of lakes) in human modified land cover classes included in the 30-year trend analyses.** The site numbers shown refer to sites that met the site inclusion requirements in Section 3.2.1 (measurements were available for at least 80% of the years and at least 80% of seasons).

Variable	Number of lakes (and sites)				
	Total	0–25%	25–50%	50–75%	75–100%
Chlorophyll <i>a</i>	20 (23)	4	6	6	4 (7)
TLI3	7 (10)	1	1	2	3 (6)
TLI4	5 (8)		1	1	3 (6)
<i>E. coli</i>	2	1			1
Secchi	19 (23)	3	6	6 (7)	4 (7)
DRP	20 (24)	3	6	7 (8)	4 (7)
TP	21 (25)	3	6	7 (8)	5 (8)

Variable	Number of lakes (and sites)				
	Total	0–25%	25–50%	50–75%	75–100%
NH <sub>4</sub> -N	20 (23)	3	6	6	5 (8)
NNN	13 (16)	1	4	4	4 (7)
TN	7 (10)	1		2	4 (7)

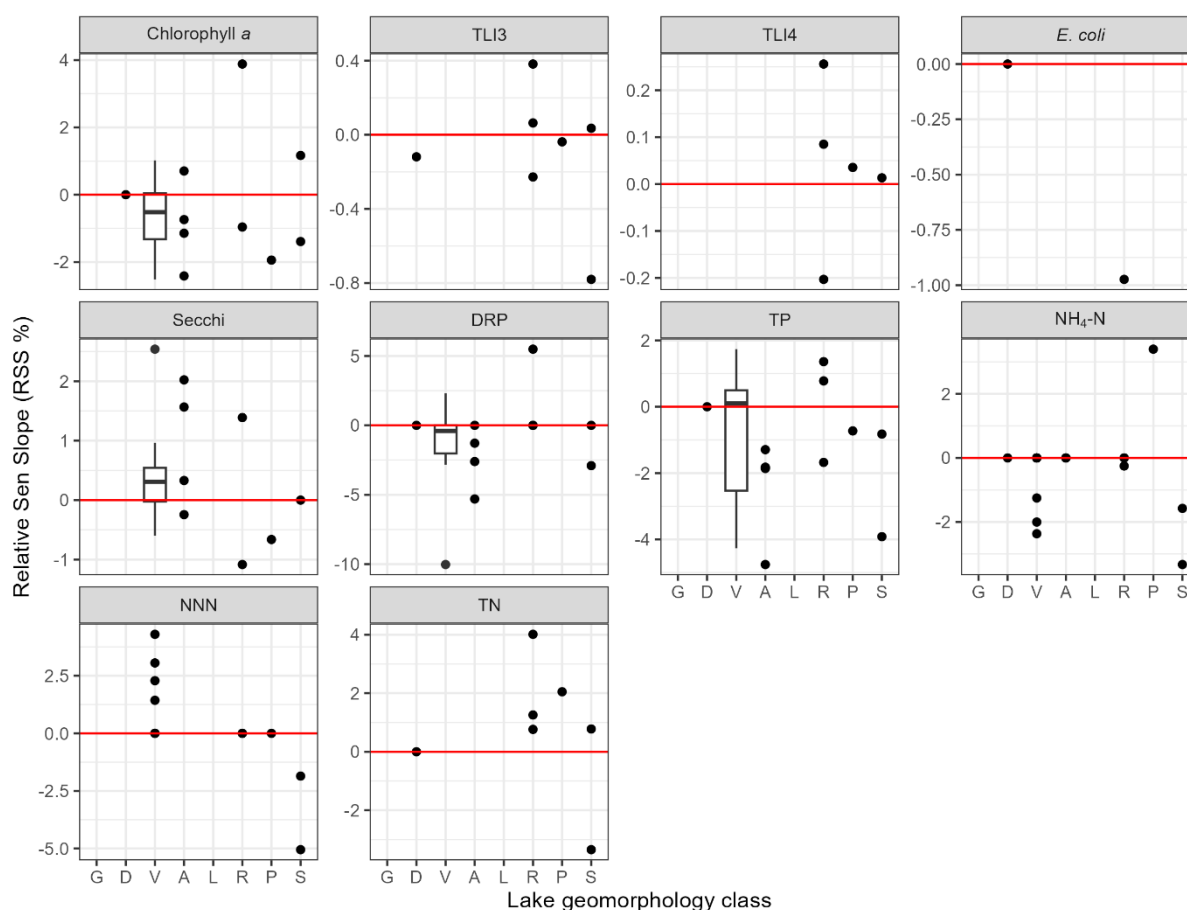


Figure 5-12: Locations of lake sites included in the 30-year trend analyses by water quality variable.

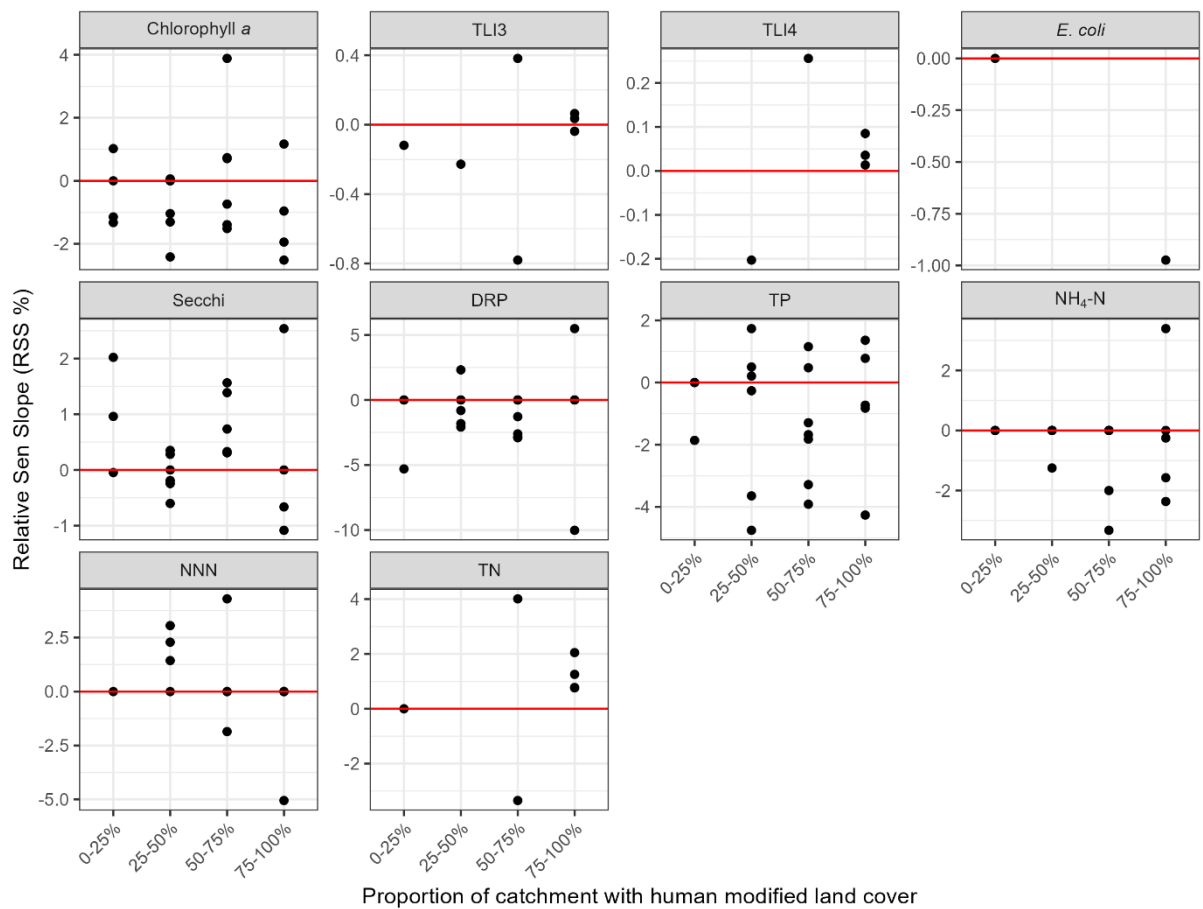
### 5.3.1 Trend rate

Box-and-whisker plots summarise the assessed trend rates for each lake and water quality variable for the 20-year period from 1995–2024 in each lake geomorphological class (Figure 5-13) and each upstream human modified land cover class (Figure 5-14). All estimated trend rates are included in these plots, irrespective of the level of confidence in the assessment (see Section 3.2.3).

The very low numbers of lakes included in the 30-year trend assessment means that the distributions shown in Figure 5-13 and Figure 5-14 are unlikely to reflect general trends in the populations of lakes in either geomorphological or upstream human modified land cover classes.



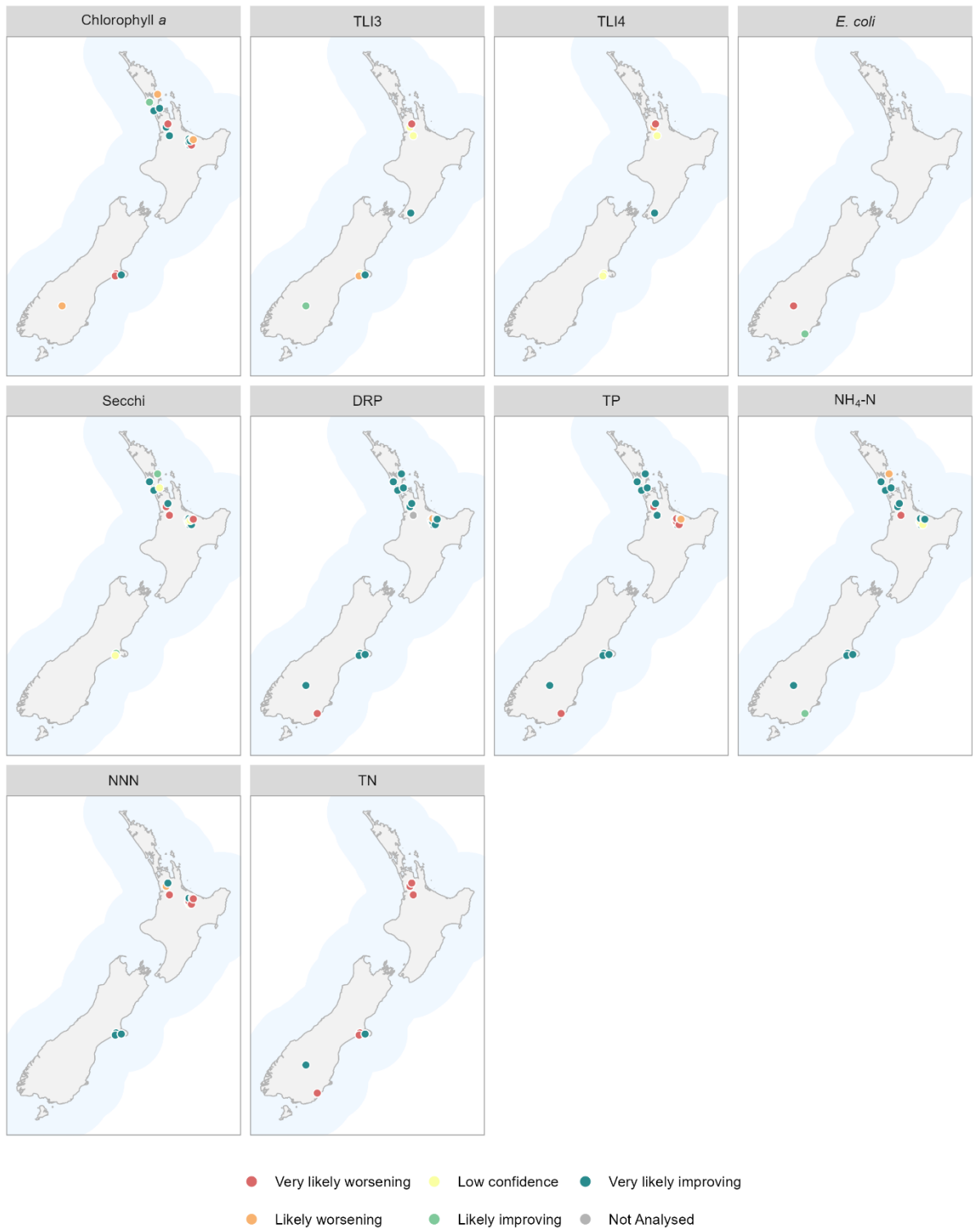
**Figure 5-13: Summary of 30-year lake trend rates in water quality variables, in lake geomorphological classes.** Geomorphological classes: glacial (G), dam (D), volcanic (V), aeolian (A), landslide (L), riverine (R), peat (P), and shoreline (S). Classes with  $\geq 10$  lakes are shown as box plots, and classes with  $< 10$  lakes are shown as points indicating individual lake values. Box plots represent percentiles of lake relative Sen slopes. Horizontal lines are the medians of lake relative Sen slopes, boxes mark interquartile ranges, and whiskers indicate  $1.5 \times$  the interquartile ranges. Lake relative Sen slopes outside of the range of whiskers are shown as black circles.



**Figure 5-14: Summary of 30-year lake trend rates in water quality variables in human modified land cover classes.** No box plots are shown as no classes had  $\geq 10$  lakes; classes with  $< 10$  lakes are shown as points indicating individual lake values.

### 5.3.2 Trend direction

The levels of confidence listed in Table 3-2 were used to categorise the confidence that 30-year trends were improving for each site  $\times$  variable combination. The spatial distributions of categorised individual sites are shown in Figure 5-15.

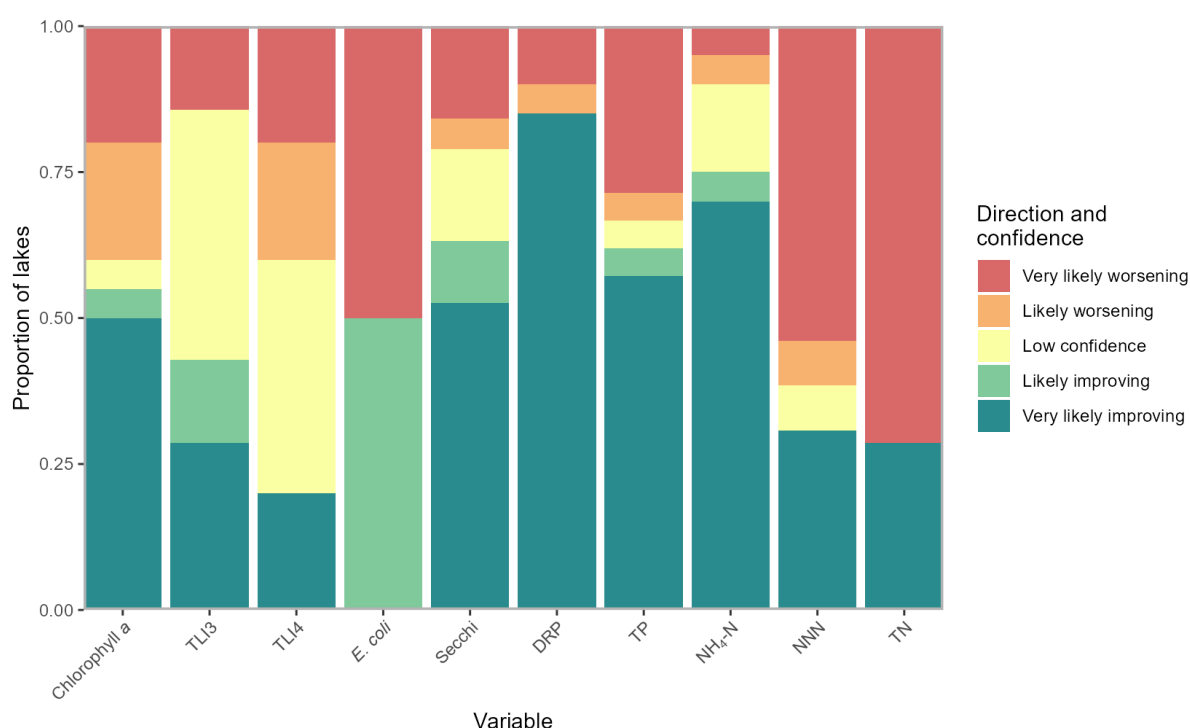


**Figure 5-15: Lake water quality monitoring sites categorised by the confidence that the 30-year trend was decreasing ( $C_i$ ) for each variable.  $C_i$  is expressed using the confidence categories in Table 3-2. Only sites that met the sampling requirements outlined in Section 3.2.1 are shown.**

### 5.3.3 Aggregate Trends

Figure 5-16 shows the proportions of lakes belonging to each of the five categorical levels of confidence for  $C_i$  defined in Table 3-2 for the 30-year trends. These plots provide a national-scale assessment of the relative proportions of lakes with improving versus worsening trends.

The national-scale proportions of improving trends ( $P_i$ ) and their confidence intervals are summarised in Table 5-9. The 30-year  $P_i$  statistics ranged from 20–85%. NNN and TN had a majority (i.e.,  $P_i < 50\%$ ) of worsening trends at the 95% confidence level. DRP, TP and  $NH_4-N$  had a majority (i.e.,  $P_i > 50\%$ ) of improving trends at the 95% confidence level. For all other variables (Chlorophyll *a*, TLI3, TLI4, *E. coli*, Secchi), the 95% confidence intervals included 50% and we could not infer generality of increases or decreases across the analysed data for these variables.



**Figure 5-16: Summary plot representing the proportion of lakes with improving 30-year time-period trends at each categorical level of confidence.** The plot shows the proportion of lakes with improving trends at levels of confidence defined in Table 3-2.

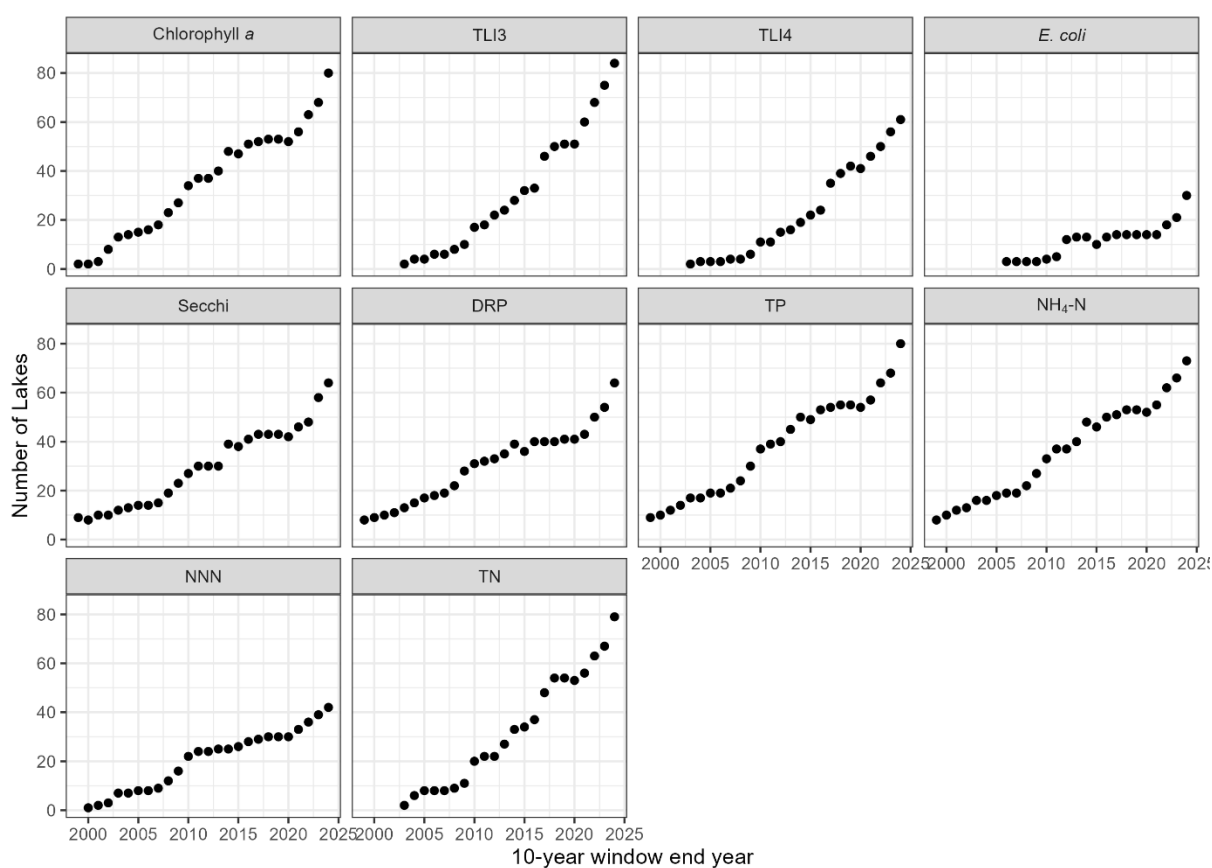
**Table 5-9: Proportions of improving trends ( $P_i$ ) for 30-year time period.**

Variable	Number of lakes	$P_i$ (%)	95% confidence interval for $P_i$ (%)
Chlorophyll <i>a</i>	20	55	43.6 – 66.4
TLI3	7	57.1	34.5 – 79.8
TLI4	5	20	0.0 – 51.5
<i>E. coli</i>	2	50	9.8 – 90.2
Secchi	19	68.4	54.6 – 82.3
DRP	20	85	79.0 – 91.0
TP	21	66.7	59.0 – 74.4
$NH_4-N$	20	85	70.3 – 99.7

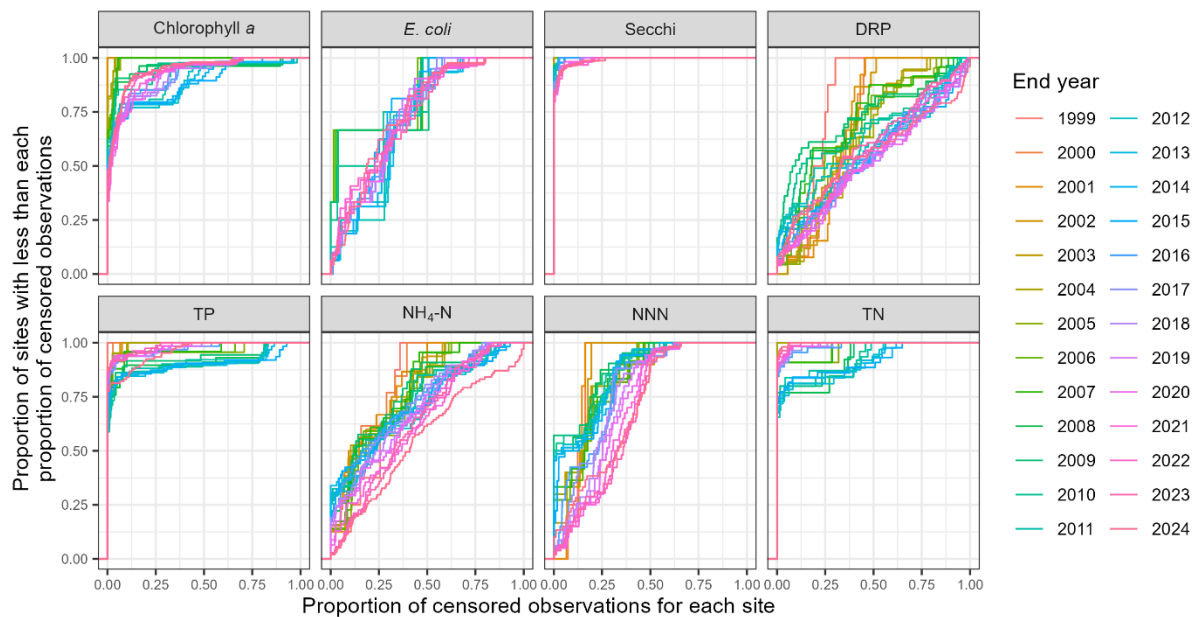
Variable	Number of lakes	P <sub>i</sub> (%)	95% confidence interval for P <sub>i</sub> (%)
NNN	13	30.8	13.8 – 47.7
TN	7	28.6	19.1 – 38.0

## 5.4 Rolling ten-year trends

The number of lakes that met the filtering rules for each of the ten-year windows and variables ranged from zero to 84. The changes in the number of sites included in the rolling ten-year trend assessment over time are shown in Figure 5-17. The effect of changes in detection limits as well as changes in sites through time were evident for all variables, to a lesser extent for Secchi, and to a greater extent for NH<sub>4</sub>-N, DRP, and NNN (Figure 5-18).



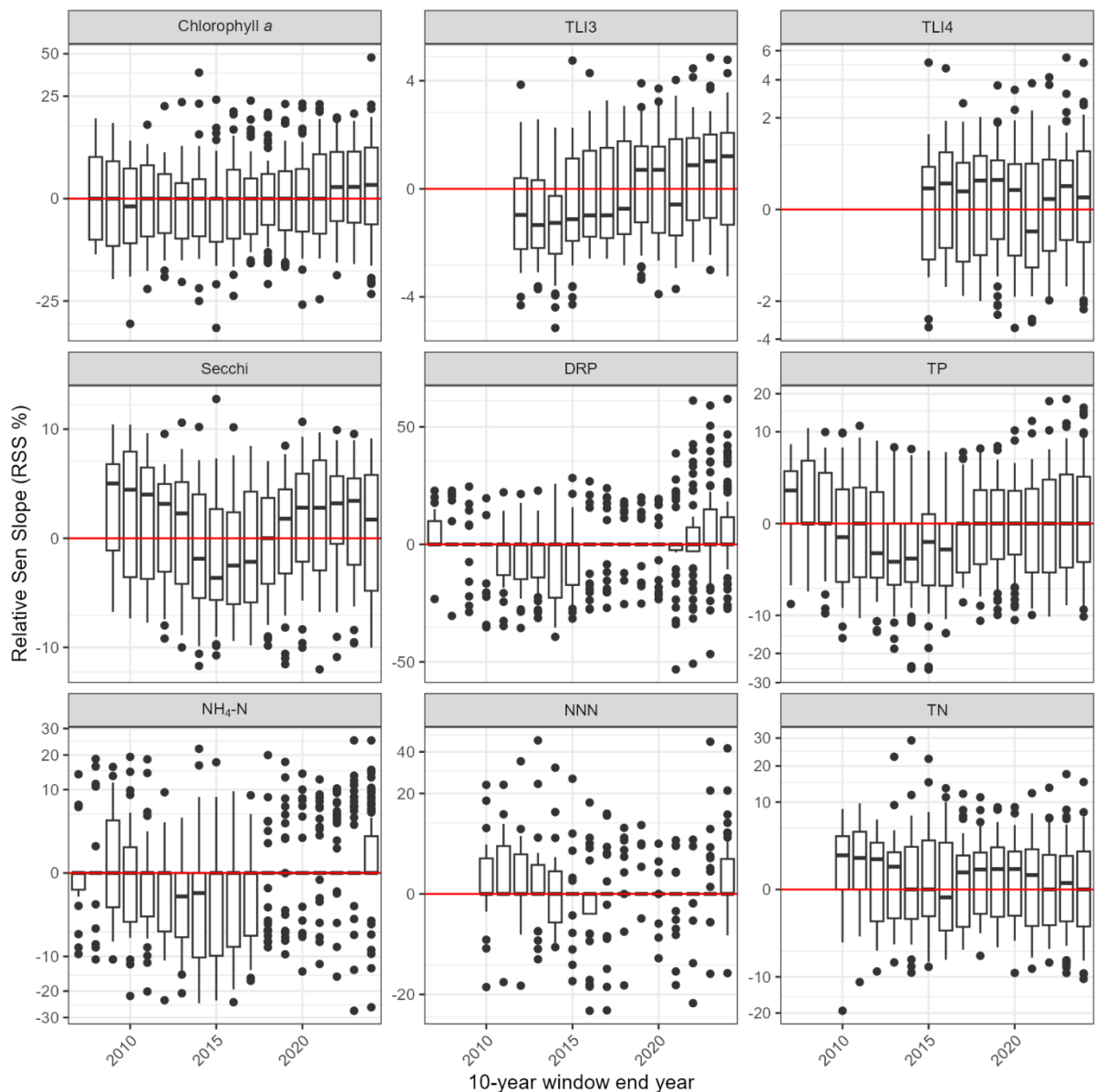
**Figure 5-17: Number of lakes included for each 10-year period in the rolling trends assessment by variable.**



**Figure 5-18: Proportion of censored data for each site × variable combination for which 10-year rolling trends were assessed.**

#### 5.4.1 Trend rate

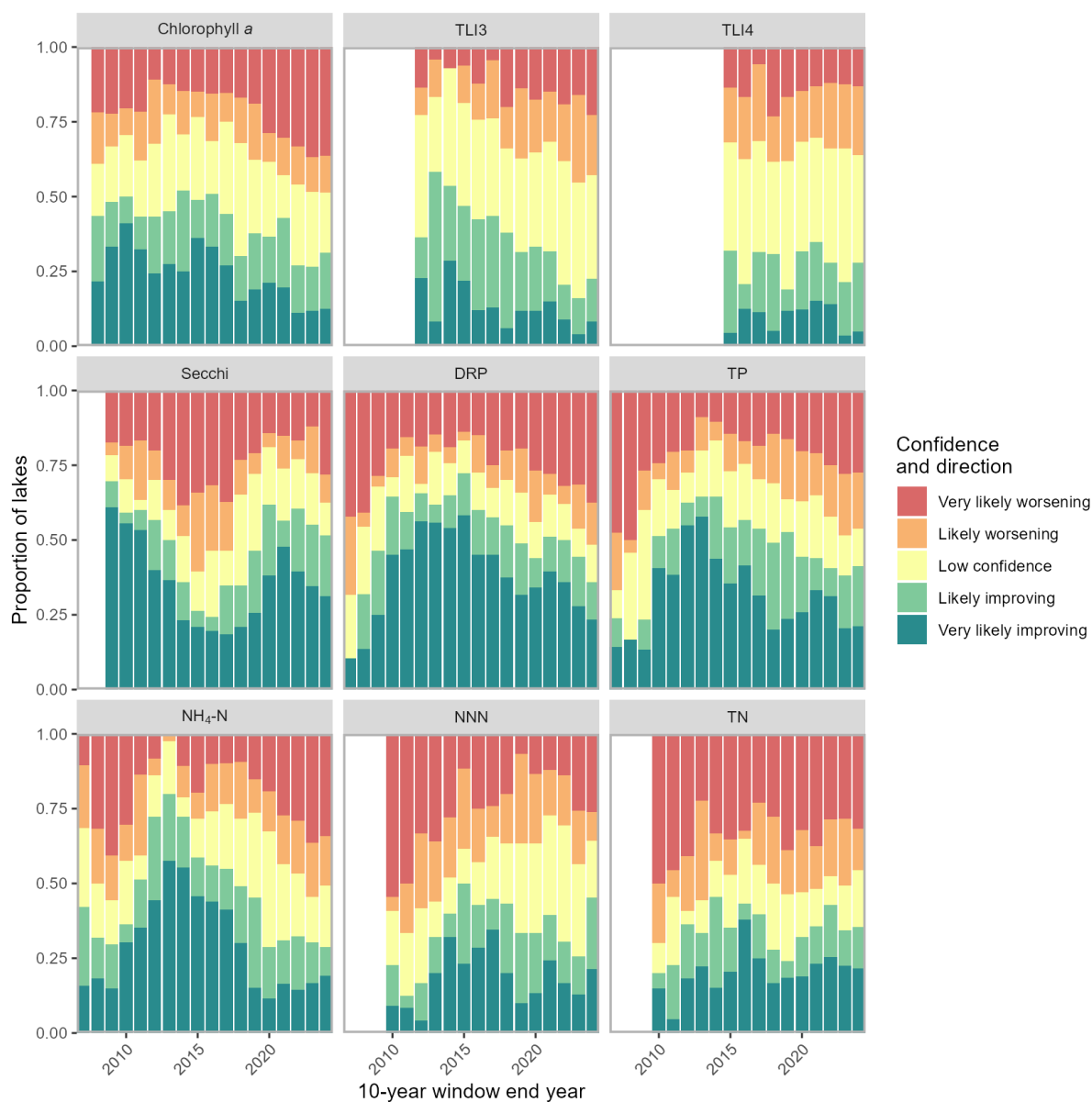
Box-and-whisker plots were used to summarise the estimated trend rates for each of the water quality variables for each 10-year window (Figure 5-19). All estimated trend rates are included in these plots provided that there were at least 20 lakes within the 10-year time window, irrespective of the levels of confidence in the assessment (see Section 3.2.3). The plots show transient patterns in median RSS for all variables. We note that the patterns may be influenced by the number of sites included in each ten-year window.



**Figure 5-19: Summary of rolling window 10-year trend rates (relative Sen slopes, %).** Box-and-whisker plots show the distributions of lake trend rates within each ten-year window. Black horizontal line in each box indicates the median of site trends, and the box indicates the inter-quartile range (IQR). Whiskers extend from the box to the largest (or smallest) values no more than  $1.5 \times \text{IQR}$  from the box. Data beyond the whiskers are shown as black circles. Note, y-axis has a signed square root transformation. The red line indicates a rate of zero. Units for each variable are given in Table 2-1.

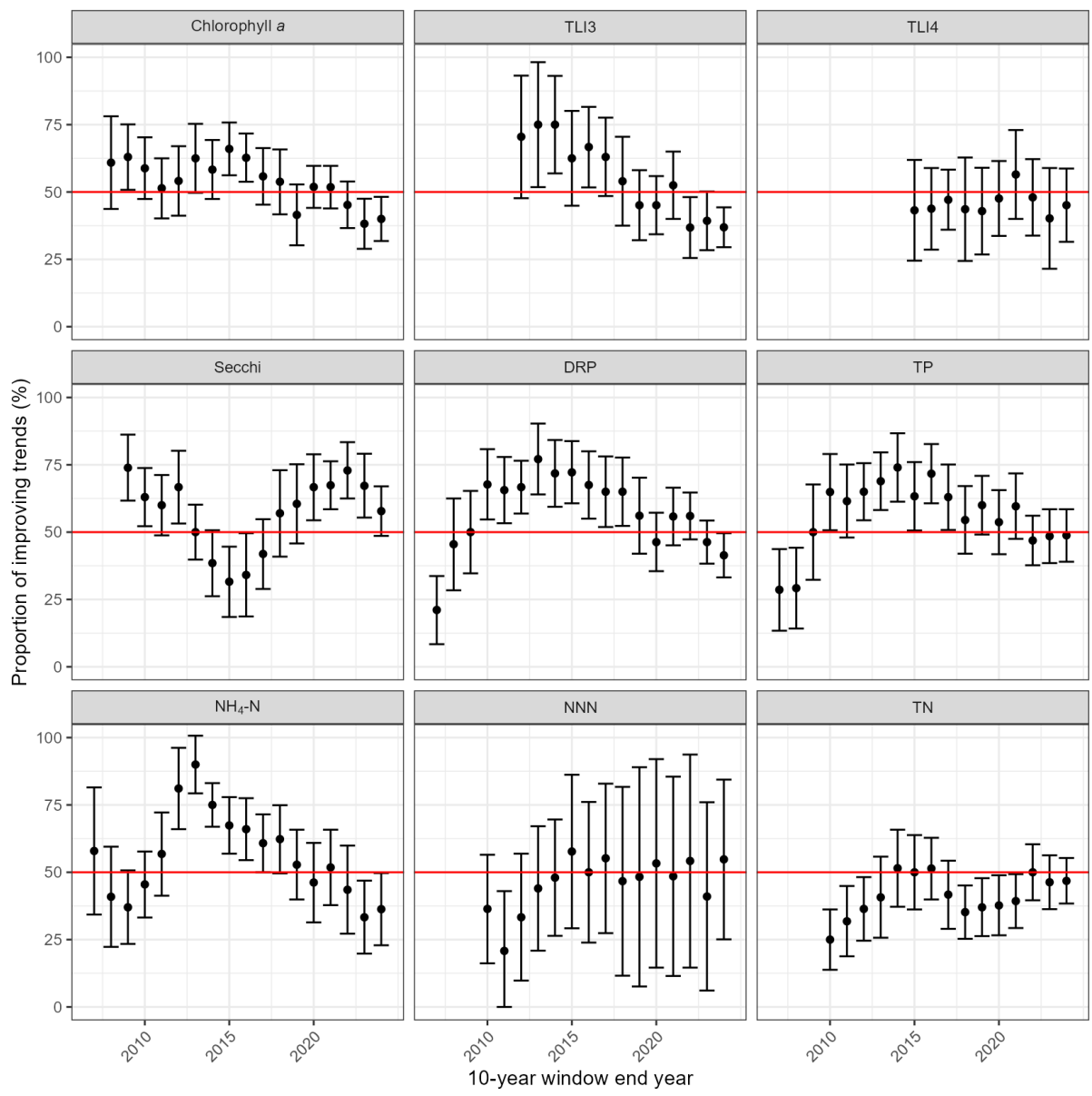
**5.4.2 Aggregate trends**

Figure 5-20 shows the proportions of lakes belonging to each of the five categorical levels of confidence for  $C_i$  defined in Table 3-2 for each of the 10-year time window trends. These plots provide a national-scale assessment of the relative proportions of lakes with improving versus worsening trends, and how these have changed through time. Time windows are only shown where there were at least 20 lakes within the 10-year time window.



**Figure 5-20: Summary plot representing the proportion of lakes with improving 10-year time-period trends at each categorical level of confidence for each time window.** The plot shows the proportion of lakes with improving trends at levels of confidence defined in Table 3-2.

The national-scale proportions of improving trends ( $P_i$ ) and their confidence intervals for each time window with more than 20 lakes are summarised in Figure 5-21. The variation in  $P_i$  varied between lake water quality variables, with the greatest variation observed for  $\text{NH}_4\text{-N}$  (ranging from 33% to 90%). TLI3, DRP, and  $\text{NH}_4\text{-N}$  all exhibited consistent reductions in the proportions of improving trends over the last 10 time windows.



**Figure 5-21: Summary plot representing the proportion of improving site trends ( $P_i$ ) for each of the rolling 10-year windows. Error bars are 95% confidence intervals for  $P_i$ .**

## 6 Discussion

The purpose of our analyses was to provide information required for national environmental reporting and to assist freshwater policy development. We overcame several challenges to fulfil this purpose by collating and analysing a large volume of data for several water quality variables collected at many lakes over a long period by a diverse set of providers.

It is impossible for state and trend results to precisely represent true, but unknown, water quality conditions with unequivocal certainty because the observed data are samples of a complete population. Furthermore, the results of our analyses will be sensitive to various features of the water quality observation data, some methodological choices, and the specifics of the monitoring network that generated the data.

We discuss factors to consider when interpreting results in Section 6.1 and describe some of the measures we applied to mitigate or better understand their effects, but emphasise that users of our results should be cognisant of these factors when interpreting results.

Several additional analyses and interpretive approaches could further enhance the utility of the information presented in this report. We acknowledge some of these in Section 6.2 but emphasise that these were outside the scope of the current work.

### 6.1 Factors to consider when interpreting results

Operational, methodological, and strategic aspects relating to the water quality data and analysis techniques have the potential to influence our results. In this section we outline examples of how these three aspects can be influential and their possible consequences for results of our state and trend analyses. First, we highlight some technical methods applied to mitigate the potential influence of operational factors on state and trend results. Second, we clarify why methodological choices of the analysis can influence state and trend results and describe how our rolling trend analyses can aid interpretation of results. Third, we describe how strategic factors can influence state and trend results but were not investigated in this report. The degree to which results are sensitive to the various influencing factors will depend on water quality variable, site, time period, and type of analysis.

#### 6.1.1 Operational

Features within water quality observation data that can arise at the operational level for a site × variable combination and have the potential to influence our results, include the following.

- The proportion and position of missing data.
- The availability of depth of sampling information to distinguish surface and sub-surface water quality observations.
- Uneven observation frequency.
- The proportion of censored data.
- Observation data precision.
- Changes in detection limits within the assessment period.
- Changes in data precision within the assessment period.

Our analysis applied transparent assumptions and robust technical procedures designed to maximise the degree to which results for a site × variable combination reflect true, but unknown, water quality conditions while accounting for the features listed above. When producing results for each site × variable combination, our approach applied the following technical procedures designed to mitigate the effects of (sometimes time-varying) features that can be contained within observation data.

- The irregular or systematic changes in sampling frequency were robustly handled by downsampling before calculating state and as part of the trend assessment.
- The censored data were robustly handled when calculating state by applying imputation based on the distribution of uncensored data.
- The censored data were robustly handled when calculating both trend rates and confidence in trend direction.
- The effect of seasonality was detected and accounted for by the trend analyses.

The proportion of censored data was relatively low at most sites for chlorophyll *a*, Secchi, TP, and TN. However, *E. coli*, DRP, NH<sub>4</sub>-N, and NNN contained considerably more censored data at more sites across all time periods compared to other variables (Figure 4-2). Therefore, the degree to which median values were influenced by data censoring varied by variable and by site. Our trend analyses produced trend rates of zero for some variables for many sites (i.e., ≥50%) due to the high proportion of censored data. A trend rate of zero indicates that any underlying trend is too small to be reliably detected given the precision of the measured observations. We note that results showing trend rates of zero are possible for any site × variable combination with a high proportion of censored values.

### 6.1.2 Methodological

Methodological choices required for water quality state and trend analysis that have the potential to influence our results include selection of analysis periods and rules for data inclusion. Estimated state is influenced by length of assessment period because of differences in sample size and an implicit assumption of stationarity of the population within an assessment period. Trend direction is specific to the trend assessment period for which it is calculated. Trend direction can change between adjacent periods as demonstrated by our analysis of rolling trends.

Our analysis applied a 5-year assessment period for state and 10-, 20-, and 30-year assessment periods for trends across all variables to maintain consistency with previous SoE reporting conducted for MfE and Stats NZ. We applied rules for inclusion within these analysis periods to ensure adequate sample sizes and seasonal coverage (where applicable) for each analysis.

We compared trends across rolling 10-year windows to aid interpretation of results for 10-year trends ending 2024. For example, transient decreases and increases in proportions of improving trends within rolling aggregate trends indicate that caution must be applied when interpreting causes and permanence of calculated trends in water quality.

When interpreting rolling trends, it should be acknowledged that there were more sites from lakes available in more recent windows as indicated by counts of lakes in different windows of the rolling trend analysis. Analysis of the proportion of censored data at each site also indicated

that this was not constant across the rolling windows for some variables. Any further analysis of the drivers of the long-term patterns should be cognisant of these features of the water quality data.

### 6.1.3 Strategic

Strategic factors relating to the water quality SoE monitoring network that have the potential to influence our results, include the following.

- Definition of lakes being applied by data providers (e.g., Waiau lagoon was supplied by Environment Southland but is not included in FENZ 2024).
- Definition of SoE monitoring being applied by data providers.
- Application of QC codes by data providers, including retrospective application of QC codes to historic data.
- The number of sites.
- Site locations being over- or under-representative of a type of lake or catchment (e.g., lowland lake, larger lake, high proportion of upstream human modified land cover) compared to other types (e.g., mountain lake, smaller lake, low proportion of upstream human modified land cover).
- Duration of monitoring at each site.
- Closing or opening of sites through time.
- Sampling frequency (e.g., monthly versus bi-monthly versus quarterly versus annually).
- Exclusion of sub-surface water quality observations.

We provided the fullest possible picture of water quality state and trends by incorporating all available SoE data that passed rules for inclusion into each analysis. This approach meant that many sites were available for calculation of 5-year states and 10-year trends for each variable for the period ending 2024. However, fewer sites were available for calculation of longer trend assessment periods. Thus, aggregate trends calculated for each variable for 10-, 20-, and 30-year trends were not commensurate because they were not calculated for the same set of sites. Arguably, comparison between aggregate trends for different periods (10-, 20-, and 30-year periods) should use a consistent set of sites in each period. Maps and counts of sites in different periods indicated that this was not the case because more sites were available in more recent times and longer running sites were biased towards regions that have maintained consistent monitoring procedures. Furthermore, maps of sites included in the different trend periods indicated that the longer the trend assessment period the less well spread, and therefore less nationally representative sites included in any analysis became.

For most variables, lakes for which state (for the period ending 2024) was calculated were reasonably well spread across a gradient of human modified land cover but not well spread geographically. There was an absence of lakes in some locations that are dominated by natural land cover such as the spine of the Southern Alps and Fiordland. Further, some councils did not provide lake monitoring data, or provided data from lake monitoring which did not meet

minimum data requirements (e.g., due to monitoring programmes established later than 1 January 2020). One prominent lake not included in the study is Lake Taupō. For Lake Taupō, SoE data provided to us included ecological water quality variables monitored at a sub-surface site (160 m deep) in open waters and *E. coli* monitored during summer at 12 sites commonly used for swimming and contact recreation during the summer (December – February) ([WRC webpage](#)). These data did not meet inclusion requirements for this study.

## 6.2 Additional analyses and interpretive approaches not applied in this report

The following additional analyses would assist interpretation of our results but were outside the scope of this report.

- We used predefined NOF thresholds to grade water quality state of each site with respect to some of the analysed water quality variables. We did not compare NOF thresholds with detection limits in the observed data. It is possible that in some cases, detection limits preclude the accurate assignment of sites to NOF grades. However, from inspecting the dataset and the results of the state assessment, we found that there were only 9 site × variable combinations (seven for chlorophyll *a* and two for NH<sub>4</sub>-N; less than 1% of all available data) where the NOF compliance statistic was below the detection level and the detection level was greater than the corresponding NOF A-band threshold.
- We relied on data providers to interpret our data request consistently (e.g., we accepted data providers' interpretation of monitoring that was carried out for SoE purposes). We did not assess the consistency with which the definition of SoE monitoring was interpreted between data providers. Differences in this definition could influence patterns of water quality state that are inferred from the results of the state analysis.
- The monitoring networks operated by each data provider were outside of our control. We did not conduct any analyses of the representativeness of the SoE monitoring networks with respect to any target population (e.g., all lake water quality within a specified period whether defined by catchment characteristics, lake characteristics, environmental classes etc).
- We did not make comparisons with any other nationwide analysis of water quality including past analyses carried out for MfE and Stats NZ for environmental reporting purposes or those associated with LAWA. We note that methodological differences between the current study and previous studies should be controlled for such a comparison to be commensurate.
- We applied linear regressions between site-medians and a gradient in human modified land cover. This approach had the advantage of being simple and transparent, but does not account for other factors that influence water quality such as geology, slope, altitude, temperature, rainfall, or type of human modification.
- Trends can be caused by various drivers including climate variability, high disturbance events such as storms/cyclones or earthquakes, and changes in land

cover/use and land management. As discussed above, trend results can also be influenced by operational factors (e.g., data completion, data precision, and detection limits), methodological choices (e.g., analysis duration, algorithm), and strategic factors (e.g., site representativeness). We did not undertake any analyses to disentangle these potential drivers of water quality trend, or attribute trend results to these drivers. We note that there is evidence that lake water quality is partly associated with climate variability (e.g., Jabbari et al. 2025).

- We used all available data to calculate rolling aggregate trends. We did not control for changes in data availability between 10-year periods.

### 6.3 Differences to previous studies

Methods used in the current report are broadly comparable with those used in the previous analyses carried out for MfE and Stats NZ for environmental reporting purposes by Whitehead et al. (2021). Differences between this and the previous study include differences in raw data availability, availability of measurement method information, availability of quality control (QC) information, improved methods for assigning observations to sampling periods, and minor updates to statistical methods. Changes in trend assessment methodology and terminology used in the report were largely made to align the reporting with recently published guidance (Helsel et al. 2020; Snelder et al. 2021; Fraser and Snelder 2021; Snelder et al. 2022). Differences between this report and Whitehead et al. (2021) are summarised below.

- As agreed with MfE and Stats NZ, five trend confidence and direction categories were used in this report instead of the nine previously applied.
- As agreed with MfE and Stats NZ, in this report we use terminology of “improving and worsening” and “proportion improving,  $P_i$ ” rather than “increasing and decreasing” and “proportion decreasing,  $P_d$ ”.
- As agreed with MfE and Stats NZ, we applied linear regressions between site-medians and a gradient in human modified land cover rather than high-intensity agricultural land cover, which was a method previously applied by Whitehead et al. (2021). Preliminary analysis (results not shown) and comparison of our results with those of Whitehead et al. (2021) indicated that regressions against human modified land cover explained more variation than regressions against specific types of land cover such high-intensity agricultural land cover. Human modified land cover was used to demonstrate how water quality state can be related to various independent variables. Data accompanying this report allow similar analysis using various independent variables depending on user needs.
- Whitehead et al. (2021) calculated state and trend at lakes for which measurements were available in at least 80% of the years (e.g., four out of five years) and at least 80% of the seasons in the state calculation period. The level of data completeness was higher for the state analysis period we analysed compared to that analysed by Whitehead et al. (2021). We therefore applied a rule that restricted site  $\times$  variable combinations in the state analyses to those with measurements for at least 90% of the sampling intervals in that period. We used a “90% rule” for state analysis to be consistent with analysis of river states (Booker

et al. 2025). We used a “80% rule” for trend analysis because there were more missing data in earlier periods and to be consistent with Whitehead et al. (2021).

- Calculation of TLI as an annual variable following Burns et al. (1999, 2000). Previous reporting calculated monthly TLI values (Whitehead et al. 2021).
- For this report we shifted observation date metadata to an adjacent sample interval when a monthly monitoring regime has two observations in a month and a missing observation in the previous/following month.
- We used an updated to account for spatial correlation in the assessed trend directions at individual sites when calculating  $P_i$  and its variance.
- We provide an overall grade for attributes with more than one numeric attribute state.
- This report uses “downsampling”, which involved selecting the observation closest to the middle of each sample interval in the state time period, rather than calculating state from all available observations.
- For trend assessment, season was here defined separately from the time increment of the assessment, rather than the season assigned to a site × variable combination being limited to the sampling interval.
- When calculating the Kendall S statistic for this report, pairs of observations that were tied in time were assigned differences of zero.
- When calculating trend rates for this report, inter-observation slopes were treated as ties and set to zero depending on the detection level and observation value within each pair, rather than applying a high censor filter.

## 6.4 Overview

In this report we present relatively straightforward summaries of state and trend results using tables and plots. Our results indicate patterns of water quality in space and time. For example, there was an overall tendency for water quality to be worse in sites with higher human modified land cover compared to sites with lower human modified land cover for all variables. However, our results also indicated complicated patterns in state and trend between water quality variables, across sites, and through time. This report is accompanied by detailed information for each lake monitoring site, such that bespoke future analysis can aggregate results for sites and their water quality conditions in various ways to meet different information requirements (e.g., grouped by region or environmental class, distributed along environmental gradients, or associated by combinations of upstream catchment characteristics).

## 7 Acknowledgements

Many thanks are due to the numerous staff across the various organisations who provided water quality data.

## 8 References

- Ballantine, D.J., Booker, D.J., Unwin, M.J., Snelder, T.H. (2010) *Analysis of national river water quality data for the period 1998–2007*. NIWA Client Report CHC2010–038 prepared for the Ministry of the Environment. NIWA, Christchurch.
- Booker, D.J., Smith, R.G., Wood, D., Fraser, C.E., Snelder, T.H. (2025) Water quality state and trends in New Zealand rivers; analysis of national data ending in 2024. *ESNZ Client Report to the Ministry for the Environment*, 95pp.
- Booker, D.J., Wilkins, M.C. (2025) Calculation of lake and lake catchment data layers for objects in the national lake database. NIWA memo to the Ministry for the Environment. June 2025, 10pp.
- Burns, N.M., Rutherford, J.C., Clayton, J.S. (1999) A monitoring and classification system for New Zealand lakes and reservoirs. *Lake and reservoir management*, 15, 255-271.
- Burns, N., Bryers, G., Bowman, E. (2000) Protocol for monitoring lake trophic levels and assessing trends in trophic state. *Client Report: 99/2 prepared for Ministry for the Environment*, 130pp.
- Camargo, J.A., Alonso, A., Salamanca, A. (2005) Nitrate toxicity to aquatic animals: a review with new data for freshwater invertebrates. *Chemosphere*, 58: 1255–1267.
- Davies-Colley, R., McBride, G. (2016) Accounting for changes in method in long-term nutrient data: recommendations based on analysis of paired SoE data from Wellington rivers. *NIWA Client Report*, HAM2016-070: 34.
- Douglas, E.M., Vogel, R.M., Kroll, C.N. (2000) Trends in floods and low flows in the United States: impact of spatial correlation. *Journal of hydrology*, 240(1-2), 90-105.
- Fraser, C., Snelder, T.H. (2021) Update to REC Land Cover categories and review of category membership rules (LWP Client Report 2021-18). LWP Ltd, Christchurch, New Zealand.
- Gehl, K. (2009) *Nitrate/nitrite toxicity*. Case studies in Environmental Medicine. Agency for Toxic Substances and Disease Registry. <http://www.atsdr.cdc.gov/csem>
- Helsel, D.R. (2005) More than obvious: better methods for interpreting nondetect data. *Environmental science & technology*, 39, 419A-423A.
- Helsel, D.R. (2012) *Statistics for Censored Environmental Data Using Minitab and R*. Second edition. Wiley.
- Helsel, D.R., Hirsch, R.M., Ryberg, K.R., Archfield, S.A., Gilroy, E.J. (2020) *Statistical methods in water resources: U.S. Geological Survey Techniques and Methods, Book 4, Chapter A3*. U.S. Geological Survey, 458 pp.
- Hickey, C.W. (2014) *Derivation of indicative ammoniacal nitrogen guidelines for the National Objectives Framework*. NIWA Memorandum MFE13504. NIWA, Hamilton.

- Hirsch, R.M., Slack, J.R., Smith, R.A. (1982) Techniques of trend analysis for monthly water quality data. *Water Resources Research*, 18: 107–121.
- Horowitz, A.J. (2013) A Review of Selected Inorganic Surface Water Quality-Monitoring Practices: Are We Really Measuring What We Think, and If So, Are We Doing It Right? *Environmental Science & Technology*, 47: 2471–2486.
- Jabbari, A., Verburg, P., Plew, D., Stewart, S.D., Zammit, C. (2025) Impacts of climate change on temperature and dissolved oxygen in an oligotrophic deep lake, Lake Taupō. *Authorea Preprints*.
- Kuczynski, A., Smith, R.G., Fraser, C.E., Larned, S.T. (2024) Environmental indicators of lake ecosystem health in Aotearoa New Zealand: Current state and trends. *Ecological Indicators*, 165, p.112185.
- Larned, S.T., Snelder, T.H., Unwin, M.J., McBride, G.B. (2016) Water quality in New Zealand rivers: current state and trends. *New Zealand Journal of Marine and Freshwater Research*, 50: 389–417.
- Larned, S.T., Snelder, T.H., Unwin, M.J., McBride, G.B., Verburg, P.H., McMillan, H.K. (2015) CHC2015-03 Prepared for Ministry for the Environment *Analysis of water quality in New Zealand lakes and rivers*. NIWA Client Report CHC2015-033 prepared for Ministry for the Environment. NIWA, Christchurch.
- Larned, S.T., Snelder, T.H., Whitehead, A.L., Fraser, C. (2018) *Water quality state and trends in New Zealand lakes: Analyses of national data ending in 2017*. NIWA Client Report 2018359CH prepared for the Ministry for the Environment. NIWA, Christchurch.
- McBride, G.B. (2005) *Using Statistical Methods for Water Quality Management: Issues, Problems and Solutions (Vol. 19)*. John Wiley & Sons.
- National Environmental Monitoring Standards (2019) National Environmental Monitoring Standards Water Quality; Part 3 of 4: Sampling, Measuring, Processing and Archiving of Discrete Lake Water Quality Data. Version: 1.0.0. March 2019. <https://www.nems.org.nz/documents/water-quality-part-3-lakes>.
- New Zealand Government (2020) *National Policy Statement for Freshwater Management 2020*. Ministry for the Environment, Wellington.
- Randall, D.J., Tsui, T.K.N. (2002) Ammonia toxicity in fish. *Marine Pollution Bulletin*, 45: 17–23.
- Snelder, T.H., Fraser, C. (2025) River water quality trends and climate variation in the Canterbury region. Including implications for assessing policy effectiveness. LWP Client Report Nos 2025–01. Christchurch, New Zealand: LWP Ltd.
- Snelder, T.H., Fraser, C., Larned, S.T., Monaghan, R., De Malmanche, S. Whitehead, A.L. (2021) Attribution of river water-quality trends to agricultural land use and climate variability in New Zealand. *Marine and Freshwater Research*, 73(1), pp.1-19.

- Snelder, T.H., Fraser, C., Whitehead, A. (2022) *Continuous Measures of Confidence in Direction of Environmental Trends at Site and Other Spatial Scales*. Environmental Challenges 9: 100601.
- Snelder, T.H., Fraser, C., Larned, S.T., Whitehead, A.L. (2021) *Guidance for the analysis of temporal trends in environmental data*. NIWA Client Report 2021017WN prepared for Envirolink (MBIE). NIWA, Christchurch.
- Verburg, P.H., Hamill, K., Unwin, M.J., Abell, J. (2010) *Lake water quality in New Zealand 2010: Status and trends*. NIWA Client report HAM2010 prepared for Ministry for the Environment. NIWA, Hamilton.
- Whitehead, A., Fraser, C., Snelder, T.H., White, R. (2021) Water quality state and trends in New Zealand lakes; Analyses of national lakes data ending in 2020. NIWA Client report 2021297CH.
- Whitehead, A., Fraser, C., Snelder, T.H., White, R., Walter, K., Woodward, S., Zammit, C. (2022) Water quality state and trends in New Zealand rivers analyses of national rivers data ending in 2020, 2021296CH: 73.
- Wood, D. (2024) Assessing and accounting for the influence of changes in laboratory measurement methods on the interpretation of long-term time-series data. *NIWA Client Report*, 2024055CH: 54.
- Yue, S., Wang, C.Y. (2002) Regional streamflow trend detection with consideration of both temporal and spatial correlation. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 22: 933-946.

## Appendix A Output data

The complete set of state analysis results is provided in the supplementary file “LakeState\_2020to2024\_251015.csv”. The metadata for each monitoring site (including human modified land cover classes) are provided in the supplementary file “LakeMetaData\_v251015.csv”. Trend results for all time periods is provided in the supplementary file “LakeTrends\_to2024\_v251015.csv”.