Ministry for the Environment
Manatū Mō Te Taiao

Making Aotearoa New Zealand the most liveable place in the world
Aotearoa – he whenua mana kura mō te tangata

Science and Technical Advisory Group:

Maintain or Improve – papers for meeting on 26 March

Paper Author	Nik Andic, Ton Snelder	Classification	**In confidence**

Meeting date	26 March 2019	Agenda item (number)	6

Paper summary:

Part 1: (Page 2) Maintain or improve: Options development with STAG sub-group.
This presents the outcome of a meeting between STAG members and officials to work through risks associated with the existing requirement to maintain or improve water quality

Part 2: (Page 7) Example of use of water quality monitoring data as information for setting objectives and assessing whether water quality has been maintained or improved
This paper provides a 'worked example' of how we can apply the STAG sub-group's approach to the use of real water quality monitoring data as information for setting objectives and assessing whether water quality has been maintained or improved.

Maintain or improve: Options development with STAG sub-group

Context

The Government has identified halting declines in water quality as a priority, and asked officials to develop a package of options to do this – the *Essential Freshwater* work programme. As part of this work programme, we are considering a range of issues with the existing National Policy Statement for Freshwater Management (the NPSFM), and how we can improve on it to achieve the Government's objective.

There are risks associated with the existing requirement to maintain or improve water quality, which mean planning could allow material declines in ecosystem health while maintaining water quality within attribute bands. The current objective/policies leave a number of questions unanswered making implementation difficult (e.g. does the requirement apply at every monitoring site? Can sites be aggregated? And how is current state determined?).

On 15 February 2019, officials met with a sub-group of STAG members to discuss these risks and co-develop technically workable options to address them. The invitation email, including an annotated agenda, is included as Appendix 1.

This paper summarises the outcome of that meeting – a technically workable approach that addresses issues with the existing requirement to maintain or improve water quality, as well as a record of outstanding issues and considerations that drove discussions.

Attached is a 'worked example' of how we might apply the sub-group's approach when setting freshwater objectives, and assessing whether water quality has been maintained, using available data.

Summary of outcome

The group understood that regional councils are faced with two tasks when implementing requirements to maintain or improve: producing a plan that is sufficient to give effect to the relevant objective and policies of the NPSFM; and assessing performance over time to determine the efficacy of plans (i.e. if water quality has actually been maintained – if not, that would indicate changes are needed).

With this in mind, the group developed the following option/approach to mitigate risks with the existing requirements:

(a) Requiring freshwater objectives to maintain the current state of water quality (as opposed to within a band); and
(b) Require councils to report on performance in terms of water quality state/achievement of these objectives alongside a wider range of information, including: pressures (e.g. human inputs and climate); higher level measures of state (e.g. overall state of ecosystem health); and responses like plan rules, methods and implementation progress.

Page 2 of this document describes key elements of the approach in more detail, and flags outstanding issues for further discussion/consideration. This should be read alongside the range of considerations that drove options development, and the notes contained in the initial email invite included on pages 3 and Appendix 1.

Key elements to be included in requirements to maintain current water quality

How do councils produce a regional plan that will maintain water quality?

1) Freshwater objectives to be set at or above *current state* and *at a site scale*, being *specific, measurable, and time-bound*

This will enable assessment of freshwater objectives against current state, as well as tracking progress towards an objective over time.

Any debate about site selection and how current state is determined to happens at the start of a planning processes and not after the fact.

Current state is determined according to A) any direction provided through Appendix 2 attributes, and B) expert input devolved to regional councils.

This is essentially the status quo. While attributes are a tool we can and do use to direct councils in determining current state, the group has mixed views on whether additional direction is needed, and how it should be provided (e.g. more specific sampling/monitoring requirements, NEMS, etc).

This is a key part of new attribute development. But we should consider how much time and resource we have to revise existing attributes, and whether there are other ways to help councils establish current state in an appropriate way. Equally, it may need to be longer term work (past 2019).

Discussions with the group also confirmed that it is technically possible to describe current state as a single point, although this is not sufficient for describing why/whether water quality has been maintained over time (hence the reporting requirement – see right).

The 'load to come'/predicted declines or improvements must be addressed via explicit timeframes for achieving objectives, and reporting on predicted state (see reporting requirements)

Outstanding issue – appropriate site selection and adequate coverage is critical

There is an opportunity for future work look at how we can support this, or establish whether intervention is necessary (e.g. regulation or funding). However, this is not something we propose to 'solve' here – we are focussing on getting the framework within the NPSFM right first.

But there may be a role for the STAG to size up this issue and give the Government a clear signal about what is needed. Or this can be assessed as part of the 2020 implementation review (required under the NPSFM).

This is the same issue as concerns that an 'FMU will be too large' – the underlying question is, do freshwater objectives (i.e. set at a site scale) apply to, and are measured against, an inappropriately large area.

Outstanding issue – whether current water quality is a fixed concept (i.e. as it was in 2014, 2019, or the date of initial planning), or something that is re-assessed over time (i.e. during future plan changes)

Currently the existing water quality is defined as the latter, something that is reassessed over time. But in parallel to this work we are considering how to address the risk this will lock in degradation over time (i.e. allow councils to aim lower following poor performance). Conversely, there are risks with a fixed concept – encouraging debate and litigation about past state, feasibility where the load to come represents a permanent decline that is already locked in under current land-use, and added complexity attributes are developed in the future.

We are continuing to work through this issue.

How do councils assess plan effectiveness/whether they have maintained water quality over time?

2) Performance is assessed and reported at regular intervals

Reporting intervals could be aligned with SOE reporting, or local government elections enabling councils to combine the task of reporting with preparing advice for incoming councillors.

This question is better suited for consultation than internal analysis, we don't necessarily have the right information or expertise to know what is best/preferred (as central government).

Reporting is *comprehensive* and tells the full story about what is happening to water quality

Discussions with the group really focussed on what information is needed to understand whether water quality has been maintained, and that simple pass/fail assessments of water quality outcomes are not sufficient to inform planning.

Reporting should include:

- Pressures – Climate influence and inputs like land-use are useful for understanding water quality outcomes (and clearly communicating uncertainty about their causes).
- State – This should include the current <u>and predicted state</u> of individual attributes (e.g. current state) AND integrate what these mean for the overall state of the values like ecosystem health AND additional measures that assist with this (e.g. biotic measures). It is important to communicate what water quality is, what it is likely to be, and what those outcomes mean for values. There is a significant overlap with ecosystem health work, which is looking at how councils can report on the overall state of ecosystem health, including potential templates for reporting, and valued metrics which could lead to additional attributes and include more integrated measures (e.g. biotic measures, etc).
- Response – Reporting should clearly communicate what action the council is taking (both regulatory and non-regulatory) to achieve freshwater objective, and their progress in implementing these – these can be useful indicators of progress towards success or failure.

Reporting should be comprehensive – include all information relevant to assessing whether water quality has been/will be maintained. <u>We can specify what this must at least include (and are looking at doing just that via ecosystem health work), and at minimum, are looking to integrate existing requirements for monitoring (incl. MCI), accounting, and reporting on implementation.</u>

There is single reporting requirement, pulling together existing and new requirements, and giving councils scope to include all relevant information

For example, plan policies may mean specific information is relevant. Policy 11.4.7 of the Canterbury Land and Water Regional Plan is to—

> *Reduce the total nitrogen load entering Te Waihora/Lake Ellesmere by restricting the losses of nitrogen from farming activities, industrial and trade processes and community sewerage systems in accordance with the target (the limit to be met over time) and limits in Tables 11(i) and 11(j).*

This implies changes in nitrogen loads, progress implementing restrictions on specific activities, and comparisons to targets and limits are all relevant. We can make sure councils refer to this in any assessment of maintaining water quality.

The NPSFM also requires regional councils to collect and make publicly available a range of information that is relevant, and should be used in any assessment of maintaining water quality, i.e. regional councils are already required to account for all takes and sources of contaminants, monitor progress towards achieving freshwater objectives, and publish progressive implementation plans.

Where reporting is not already required (e.g. predicted future state, and the influence of climate) we can direct councils/give them scope to include this as well, and tell a fuller story.

Considerations that drove options development, or came up in discussions:

- **Limited ability to link changes in water quality to causes**: We have a limited ability to explain changes in water quality. Simply assessing performance in terms of water quality at points A and B will not provide enough information to attribute that to a cause and determine whether the plan was effective (or whether something else like climate was at play). There is a large amount of uncertainty here that needs to be communicated, and more information is needed to tell the full story.

- **Inputs are important too**: A narrow focus on water quality outcomes is unnecessarily limiting – inputs (e.g. changes in land use) can also provide useful information and tell a fuller story.

- **Taking a bottom-up approach is preferable**: It's more practical to think about maintaining water quality by starting at the individual site level, and building up a picture for the catchment from there (as opposed to starting at the catchment level and somehow deriving what needs to happen at the site level).

- **Adequate monitoring sites and coverage is critical**.

- **Load to come**: Need to consider how unavoidable or predicted declines are accounted for.

- **Implications for allocation and trading**: How requirements to maintain water quality are expressed (e.g. as maintaining current concentrations of a contaminant at every site) has implications for allocation systems and trade-ability of discharge rights.

- **The NOF is incomplete**: We are considering what 'maintaining' water quality means without a 'complete' set of attributes that need to be managed.
  - Other measures may be complex or have peculiarities that mean whatever approach we take is not appropriate (e.g. requiring specific monitoring periods/approaches, have complex relationships with other attributes, etc.).
  - Having adequate measures of water quality is critical to knowing whether you are maintaining in a meaningful way (i.e. how ecosystems are actually doing).
  - Measures that are yet to be included could provide a better way to assess maintenance (e.g. measures that integrate multiple aspects of ecosystem health, such as fish, which are less noisy).

- '**Overall' is confusing**: Still unclear what, if anything, the word 'overall' means within Objective A2. (Nb: it is very likely any option will involve removing this, in line with previous advice to Ministers).

Appendix 1: Initial email invite and notes contained within

-----Original Appointment-----
From: Nik Andic
Sent: Thursday, 7 February 2019 2:29 PM
To: Nik Andic; Ton Snelder; bryce.cooper (bryce.cooper@niwa.co.nz); acanning (acanning@fishandgame.org.nz); Clive Howard-Williams (niwa.co.nz); Jon Roygard (jon.roygard@horizons.govt.nz); mike.joy (mike.joy@vuw.ac.nz); Jennifer Price; Vicki Addison; Carl Howarth; Stephen Fragaszy; Jo Burton
Subject: Essential Freshwater: Invitation to a STAG sub-group on the requirement to 'maintain or improve'
When: Friday, 15 February 2019 10:00 AM-4:00 PM (UTC+12:00) Auckland, Wellington.
Where: Ministry for the Environment, Room 3A

Hi all,

Thank you again for agreeing to participate in this sub-group looking at the requirement to maintain or improve.

Based on people's availability it looks like Friday 15 Feb 2019 is the best option, but there is still time to change if it going to be a problem for you – please respond ASAP if that's the case.

As discussed, we will use a worked example to develop approaches for maintaining. We will have a large screen in the room, and will be able to look at water quality data, graph it, map it, etc. I suggest we use an area with a reasonably developed plan such as Canterbury (see plan here). When reporting back to the STAG on a preferred approach, we will try and explain it in a real world scenario (e.g. in the Hinds catchment it would look like this).

I have booked us a room for the day, and suggest the following agenda. Also, rather than send out a separate paper, we have included some suggested objectives, criteria, and questions for the day further below.

Agenda

10:00 to 11:00 – Triaging the issues. What concerns you about current settings? Is it a lack of clarity, an implementation issue, or a fundamental issue with the policy (or something else)? We propose to record these and try to address them by the end of the day. We will focus the day on fundamental issues with the policy, although we can pick up implementation issue and others at another meeting if there is interest.

11:00 to 12:00 – Agree to criteria. We have suggested some criteria that are important from our point of view. But there may be others that are important from yours. Equally we may have suggested something that is problematic – we are keen to form a joint view with the group on what criteria should be applied and why.

12:00 to 13:00 – Lunch. We will provide food.

13:00 to 14:30 – Mock-up approaches. We want to develop options for 'maintaining' water quality that would address our concerns (i.e. potential for water quality to decline, ability for councils to implement this in planning).

14:30 to 15:30 – Decision/recommendation time. We want to assess options against our criteria, and if possible, decide on a preferred approach or narrow down the list of option. The next step would be to report back to the wider STAG on this, and our reasons.

Objectives

The purpose of this workshop is to develop an approach to 'maintaining' water quality that:

a) Addresses the risk of material declines in ecosystem health if maintaining within attribute bands (e.g. by adding/modifying band thresholds, or maintaining at current state); and

b) Provides regional councils with sufficient detail to implement the requirement through planning (e.g. how to determine current state, what level of change constitutes improvement or decline, etc).

Suggested criteria

The following criteria are a starting point for assessing possible approaches. Please feel free to suggest additional criteria, or why any of these shouldn't apply!

- **Unambiguous** – Uncertainty and low levels of confidence will be a significant issue for regional councils implementing the policy. As far as possible, how we characterise maintaining current state should minimise uncertainty and scope for debate.
- **Suitability for regional planning** – The ultimate goal of this exercise is to inform possible amendments to the NPSFM, which in turn must be implemented by regional councils, through planning. Planners are really faced with two tasks when giving effect to Objective A2 and related policies:
    a. *producing a plan that is sufficient* to give effect to it (i.e. freshwater objectives aim high enough), and
    b. *assessing performance over time* to determine the efficacy of plans (i.e. has water quality actually been maintained – if not, that would indicate changes are needed upon plan review or earlier).

  It is not about individual compliance or pollution incidents.
- **Coherence with the wider monitoring and reporting** – As far as possible we should characterise 'maintaining current state' in a way that works with other monitoring and reporting practices/requirements (e.g. Environmental Reporting, state of the environment reporting, the specifics of current attributes, etc).
- **Feasibility within work programme timeframes/resources**.

Questions for discussion

1. **How should we determine the 'current state' of water quality?** Is this a 5 or 10 year average? Something else?
2. **Do the current attributes, or technical constrains tell us anything about how 'current state' should be determined?** For example, given the actual frequency of monitoring, do 'annual medians' and '80^{th} percentiles' etc actually require X years of data get these with sufficient confidence?
3. **How should multiple monitoring sites be treated?** Should sites be treated separately when setting objectives and monitoring progress? Is it defensible to aggregate these, and if so, how? Does it matter as long as plans are explicit and can be tested on a case by case basis?
4. **How can we account for noise,** and distinguish this from improvement/maintenance/decline? What level of change can actually be detected (e.g. given actual monitoring frequency and how we determine current state)?
5. **What is an appropriate sensitivity?** What level of change is meaningful?
6. **What level of specificity does the NPSFM need to contain**, and what is more appropriate in technical guidance?

Example of use of water quality monitoring data as information for setting objectives and assessing whether water quality has been maintained or improved

Introduction

This paper provides a 'worked example' of how we can apply the STAG sub-group's approach to the use of real water quality monitoring data as information for setting objectives and assessing whether water quality has been maintained or improved. The example has the following aims:

(a) demonstrate the current state of water quality at site-scale and consider how this information can inform more specific freshwater objectives than NOF bands; and
(b) explore whether we can use state and trend information to assess whether water quality has been maintained or improved.

The example is intended to highlight a range of issues with the above, and prompt questions for discussion. It does not necessarily provide answers to all questions, particularly where this would involve making value judgments and is not a scientific decision. Discussions can, and should be broader the issues and questions included in this document.

The example is based on water quality data for; dissolved inorganic nitrogen (DIN) nitrate-nitrogen (NO_3-N and periphyton biomass as chlorophyll *a* (CHLA), for one FMU in the Manawatū-Wanganui region; the Manawatū River catchment.

The One Plan predates the NPS-FM, and this example is not intended to comment on the extent to which the One Plan gives effect to the NPS-FM. In addition, the One Plan's provisions are not necessarily configured in the same way as a plan that is developed to give effect to the NPS-FM. For this example, the One Plan's targets for CHLA are treated as objectives and DIN is treated as a mechanism to achieve the CHLA objective and are referred to as DIN targets. In this example, we report state and trend analyses for both CHLA and DIN and treat them as equally important. However, we note that under the NPS-FM, CHLA would be the objective and, in practice, this may have implications for assessing whether water quality has been maintained (e.g., where CHLA is unchanged despite decreases/increases in DIN concentrations).

The NPS-FM also has an attribute for NO_3-N, which is the dominant form of nitrogen in DIN. The NPS-FM NO_3-N attribute is associated with avoiding toxicity, which generally occurs at concentrations of NO_3-N that are higher than those set for managing trophic state (i.e., periphyton biomass in rivers). The One Plan has set DIN criteria for managing trophic state (i.e., periphyton) in all rivers that are less than the NOF A/B band threshold for NO_3-N toxicity and therefore the NO_3-N toxicity attribute can be considered redundant in the One Plan. However, this example has investigated the current state of NO_3-N in comparison to the NOF bands in order to compare the current state with the width of a NOF band. To avoid complicating this example, this report has focused on the state of CHLA at sites in the FMU relative to the relevant NOF bands rather than comparing them with the One Plan objectives.

Data

The input data are derived from Horizons Regional Council's water quality monitoring programme. A complete description of the programme, datasets and results for all variables are available in Fraser and Snelder (2018).

The concentrations of DIN NO_3-N and CHLA are measured at state of environment monitoring sites (i.e., representative of general conditions rather than downstream of point sources) each month. For the five-year period ending 2017, there were 56 state of environment monitoring sites in the Manawatū River FMU with at least 30 observations of DIN and NO_3-N and there were 34 sites with at least 27 CHLA observations. In the following sections, assessments of state and trends in DIN NO_3-N and CHLA that were made from these data are presented.

State

The National Objectives Framework (NOF) bands for NO_3-N toxicity attribute are defined by two statistics calculated from monitoring observations; the median and the 95th percentile (Table 1). A total of 56 state of environment sites in the Manawatū FMU had between 30 and 54 NO_3-N observations in the 5-year period ending 2017).

The median and 95th percentiles of NO_3-N observations were calculated for each site and the 95% confidence intervals for both statistics were estimated using boot strap analysis. Sites were allocated to the A or B band based on whichever statistic produced the lower band (Figure 1).

Table 1. NOF bands for NO_3-N toxicity attribute. The units of the numeric attributes are mg L^{-1}.

Attribute band	Numeric attribute state	
	Annual Median	**Annual 95th percentile**
A	≤1.0	≤1.5
B	>1.0 and ≤2.4	>1.5 and ≤3.5
C	>2.4 and ≤6.9	>3.5 and ≤9.8
D	>6.9	>9.8

Sites in the FMU were in either the A (44 sites) or B (12 sites) NO_3-N toxicity bands (Figure 1). The 95% confidence intervals for the median and the 95th percentile NO_3-N concentrations were generally much smaller than the width of the A and B bands (Figure 1). It is noted that some sites had 95% confidence intervals that crossed the band thresholds. This means that it was uncertain at the 95% level of confidence what band the site belongs to. This point is considered in more detail below for DIN.

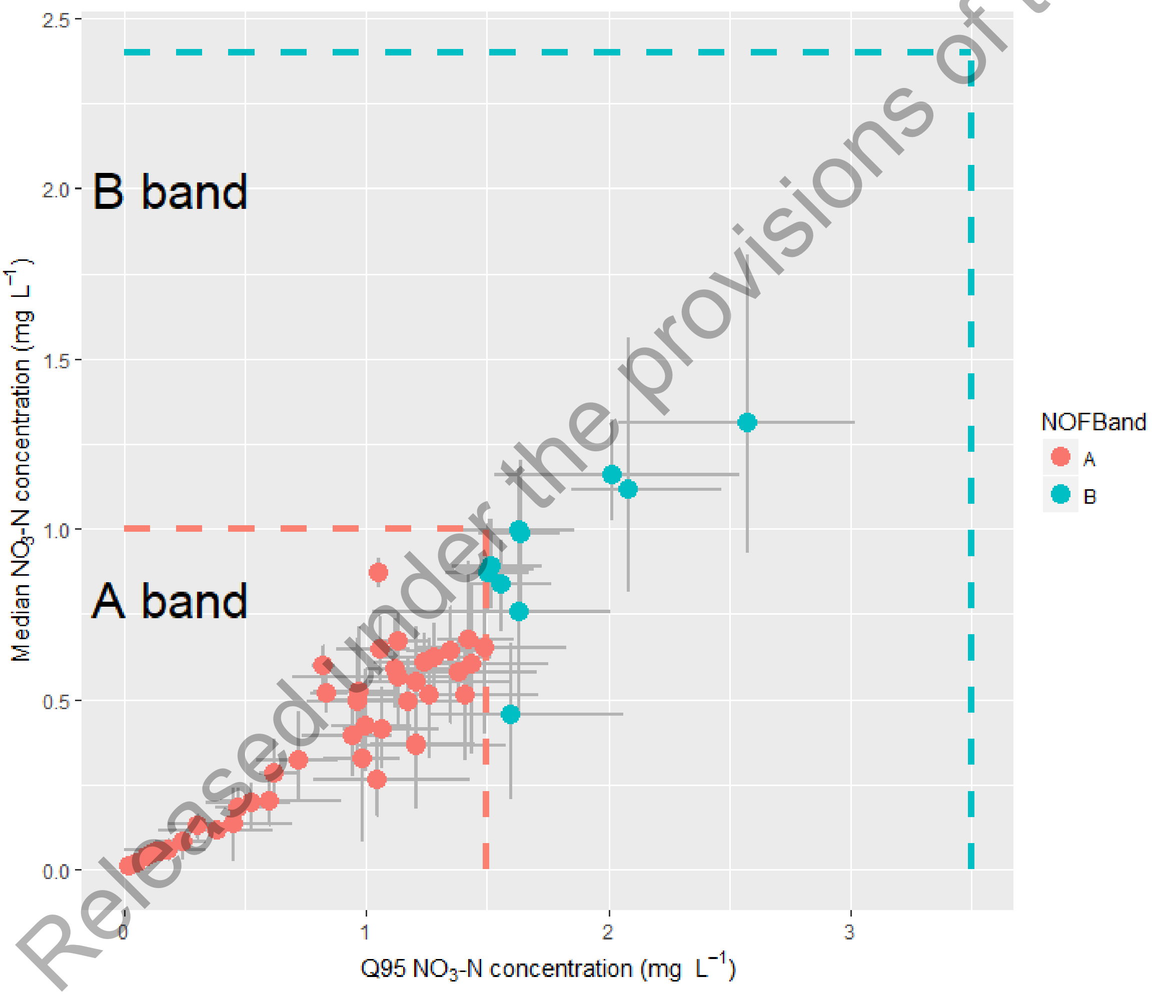

Figure 1. State of 56 monitoring sites in the FMU with respect to the NOF nitrate-nitrogen toxicity attribute bands. The points indicate the 95th percentile and median values of nitrate-nitrogen at each site. The horizontal and vertical error bars indicate the 95% confidence intervals of the 95th percentile and median values respectively. The dotted lines indicate the NOF A/B and B/C band thresholds and the points indicate the band assignment.

The One Plan sets targets for DIN state based on the mean value of the observations made when the river flow is below the 20th flow exceedance percentile (Table 2). The DIN targets are spatially variable within the FMU with more stringent (i.e., lower mean concentrations) applying at headwater locations and less stringent objectives in main stem sites (Figure 2). The One Plan sets a minimum sample size of 30 observations. In this study, the observations of DIN in the 5-year period ending 2017 were used to calculate the mean.

Table 2. Specification of the DIN targets in the Horizons One Plan. The value 'X' differs by site and in the Manawatū FMU can be one of four values 0.07, 0.11, 0.167 and 0.444 (see Figure 2.

Target name	Method[1]	Flow percentile	Sample size required	Target description
DIN	Mean	80	30	The annual average concentration of soluble inorganic nitrogen (DIN) when the river flow is at or below the 20th flow exceedance percentile must not exceed X grams per cubic metre, unless natural levels already exceed this.

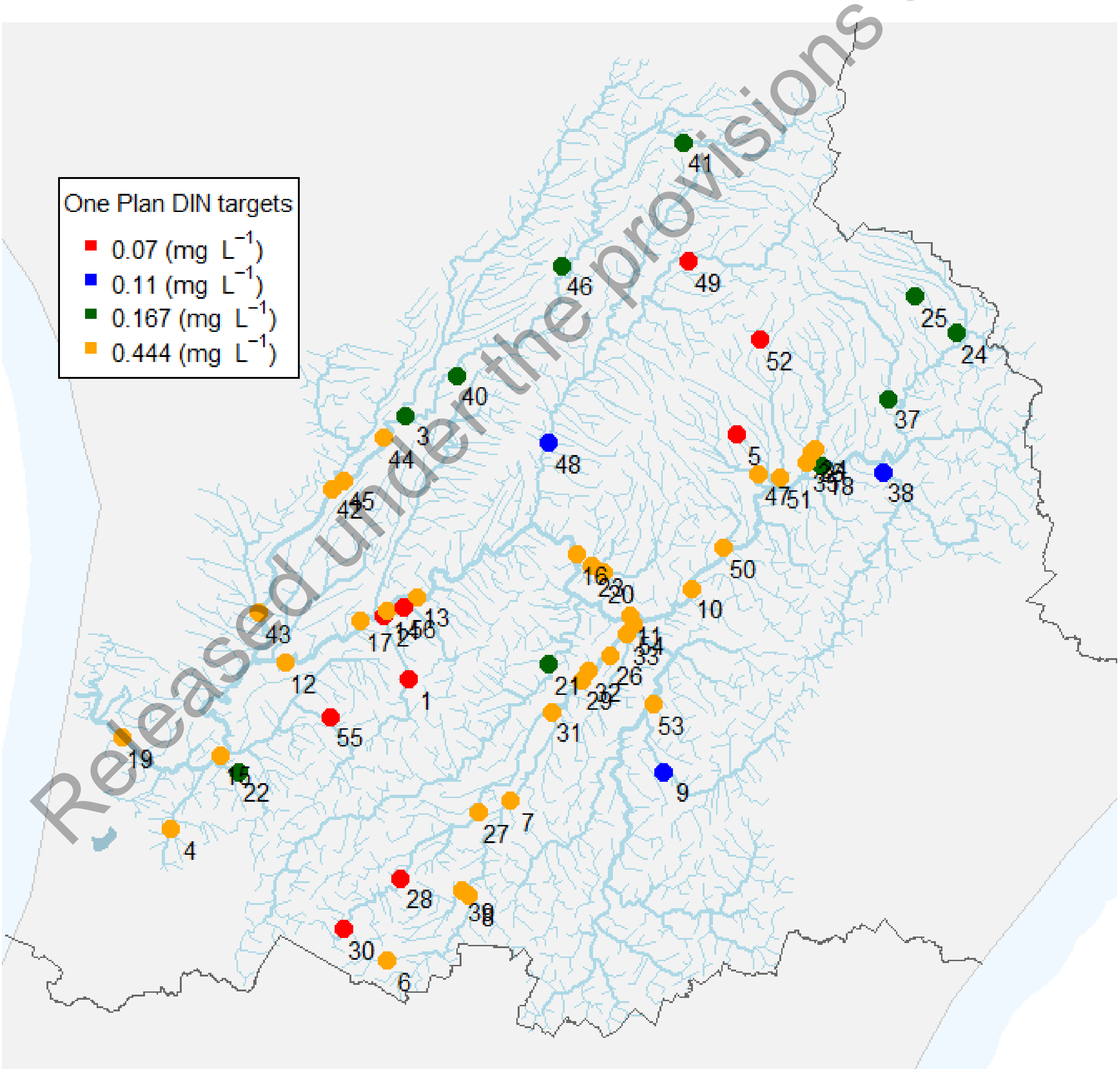

Figure 2. One Plan DIN targets at SoE Sites. The sites are numbered so that their current DIN can be seen on Figure 4.

There were a total of 56 state of environment sites in the Manawatū FMU where DIN had been observed and there were 46 sites with between 30 and 54 observations at flows at or below the 20th exceedance percentile in the 5-year period (i.e., 10 sites had insufficient data for state to be assessed, Figure 4).

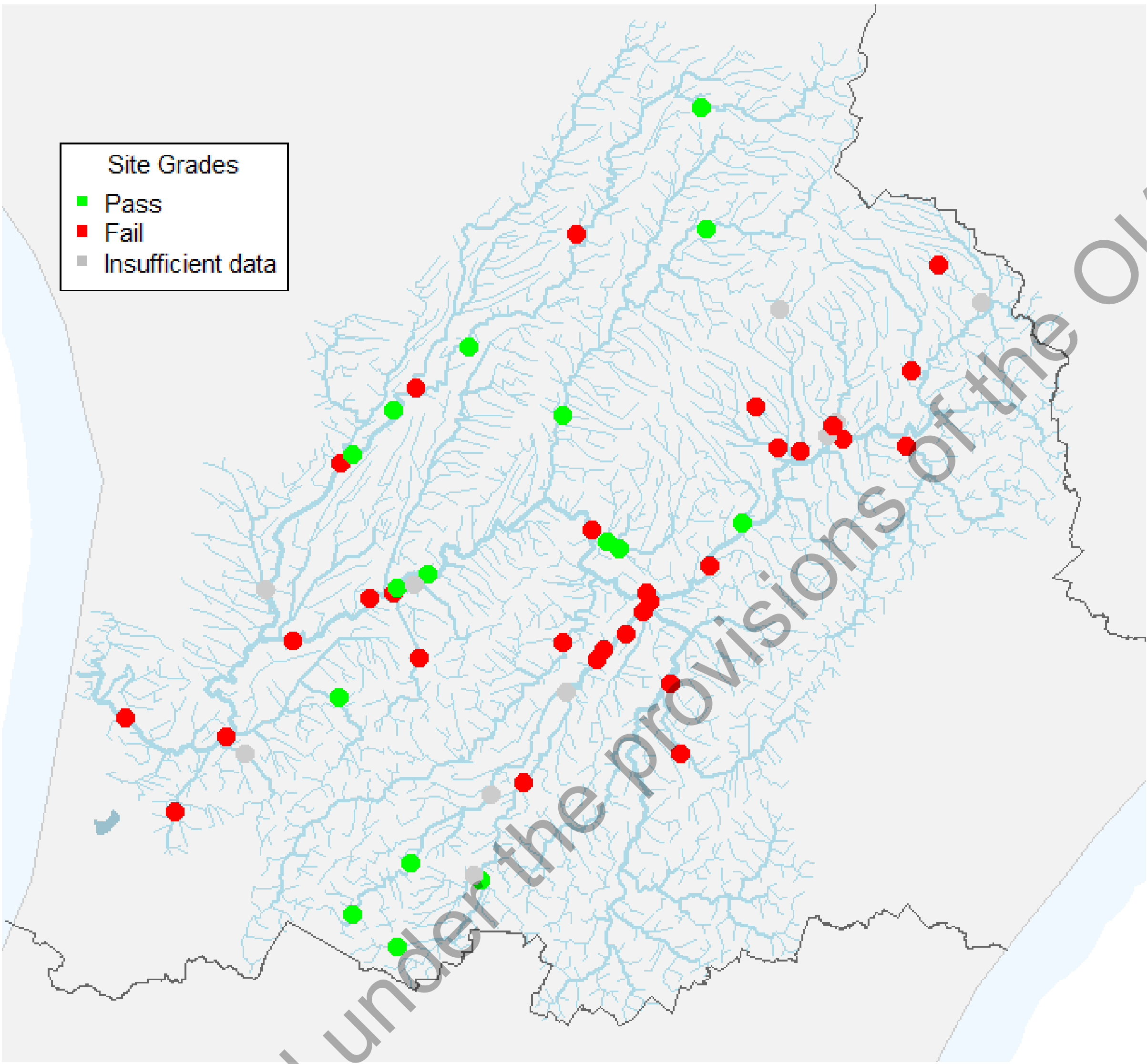

Figure 3. Map of the Manawatū River catchment DIN monitoring site locations with state of environment sites graded according to DIN concentration compared to the One Plan targets.

Sites were assigned to a pass or fail states based on the mean value compared to the target (Figure 2). Based purely on the mean values (i.e., ignoring the precision of the estimated mean), there were 16 passing sites and 30 failing (Figure 3). However, there were 12 sites that could not be classified pass/fail (at the 95% level of confidence) because the 95% confidence interval for their mean included their DIN concentration target (Figure 4).

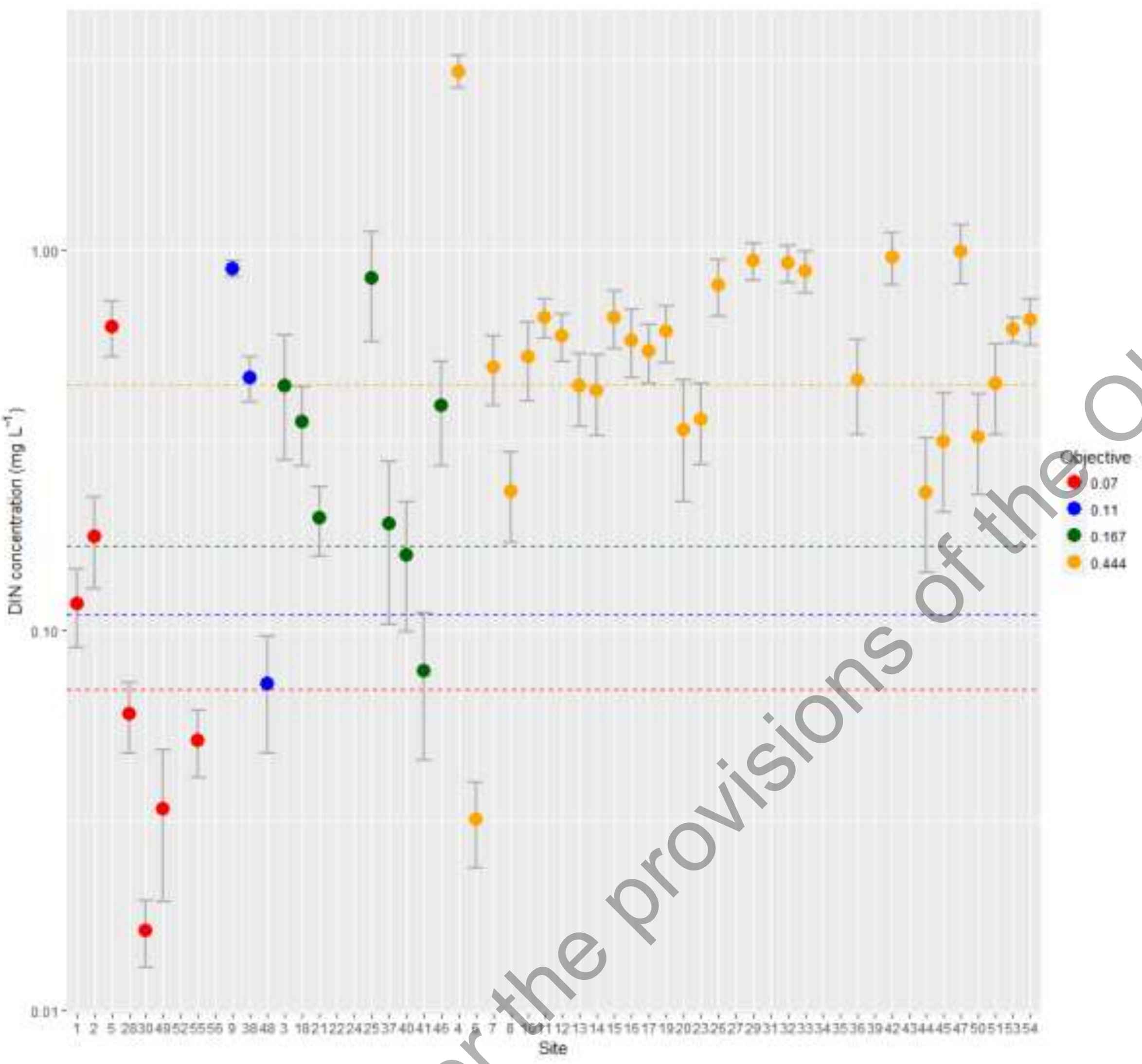

Figure 4. Evaluated state (mean concentration of DIN) at each site compared to the site's target. Each point represents a site, which is coloured by its DIN target. The error bars indicate the 95% confidence interval of the mean at each site. The horizontal lines indicate the DIN concentrations corresponding to the four targets. Sites whose 95% confidence interval for the mean value includes the DIN concentration target have an equivocal site grading (i.e. we are unconfident whether they pass or fail). See Figure 2 for site locations.

The National Objectives Framework (NOF) bands for the CHLA attribute are defined by the 92nd percentile value calculated from monitoring observations (Table 3). A total of 34 state of environment sites in the Manawatū FMU had between 27 and 71 CHLA observations in the 5-year period ending 2017).

Table 3. NOF bands for the CHLA attribute. The units of the CHLA numeric attribute is mg m^{-3}.

Attribute band	92nd percentile value	Annual 95th percentile
A	<50	≤1.5
B	50 - 120	>1.5 and ≤3.5
C	120 - 200	>3.5 and ≤9.8
D	> 200	>9.8

The 92nd percentiles of CHLA observations were calculated for each site. For simplicity the 95% confidence intervals for these statistics are not presented. Sites were allocated to the NOF attribute band based the thresholds shown in Table 3 (Figure 1). Based purely on the 92nd percentile values (i.e., ignoring the precision of the estimated statistic), there were 17, 8, 6 and 3 sites in the A, B, C and D bands respectively (Figure 5). However, as for DIN, the grading of some sites will be equivocal (at the 95% level of confidence) because the 95% confidence interval for the summary statistic will include a band threshold.

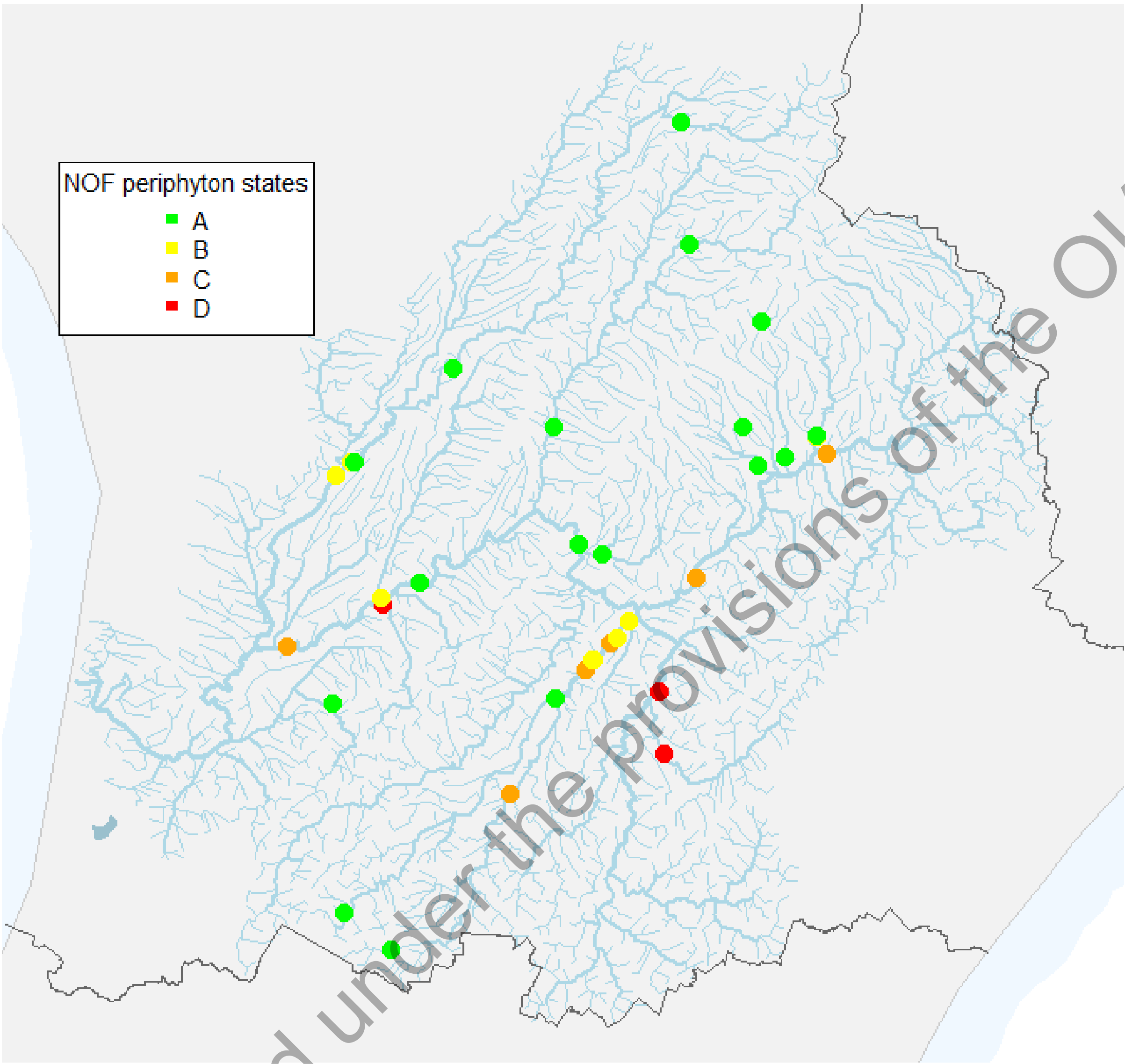

Figure 5. Map of the Manawatū River catchment periphyton monitoring site locations with state of environment sites graded according to the NOF attribute bands.

The One Plan DIN target can be compared to the NOF band grades for CHLA at 29 state of environment monitoring sites. Of the 14 comparison sites that were in the NOF A grade for CHLA, four failed the One Plan DIN target and three of the four sites in the NOF B grade for CHLA failed the One Plan DIN target (Table 4). Therefore, there is a degree of ambiguity about the achievement of objectives in that, despite the failure to achieve the DIN target, the sites achieved good states (A and B NOF bands) for CHLA.

Table 4. Comparison of site state based on One Plan DIN target grade and NOF CHLA objective grade.

One Plan DIN target grade	NOF CHLA objective grade			
	A	B	C	D
Fail	4	3	6	0
Pass	10	1	0	0

Trends

Trends in DIN were evaluated for two time periods; the 5 and 10-year periods ending 2017. The results provided here are raw trends (i.e., not flow adjusted) because flows were not available for some sites. Fraser and Snelder (2018) showed that there are some differences between raw and flow adjusted trends for some sites, but that conclusions about general water quality changes (i.e., over all sites) were similar for raw and flow adjusted trends.

For the 5-year period, there were 49 sites with sufficient data for trend analysis of DIN. Of these, 32 sites (65%) had degrading trends established at the 95% level of confidence and one site has an improving trend (Figure 6). All other (16) sites had insufficient data to establish trend direction at the 95% level of confidence. However, the aggregate trend for the whole FMU, as indicated by the proportion of improving trends PIT) indicated that 11% of sites were improving (standard error of 3.1%). This indicates that 89% of sites were degrading.

For the 10-year period, there were 40 sites with sufficient data for trend analysis. Of these, one site with a degrading trend established at the 95% level of confidence and six (15%) with improving trends (Figure 7). All other (33 or 83%) sites had insufficient data to establish trend direction at the 95% level of confidence. The aggregate trend for the whole FMU, as indicated by the proportion of improving trends PIT) indicated that 71% of sites were improving (standard error of 5.9%). Note that this implies that 29% of sites were degrading over the 10-year period.

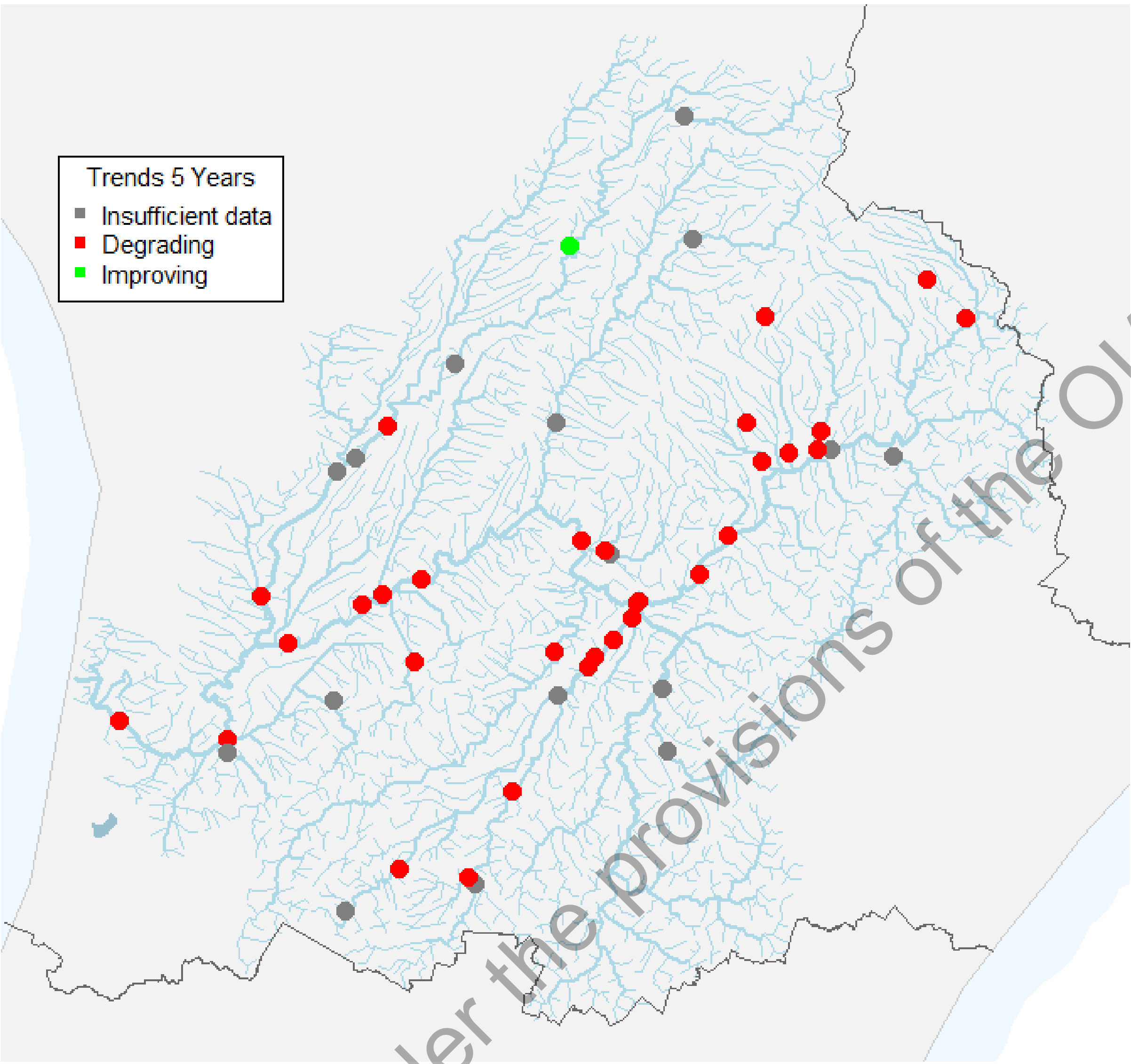

Figure 6. Map of the Manawatū River catchment periphyton monitoring site locations with sites coloured by 5-year trend evaluations.

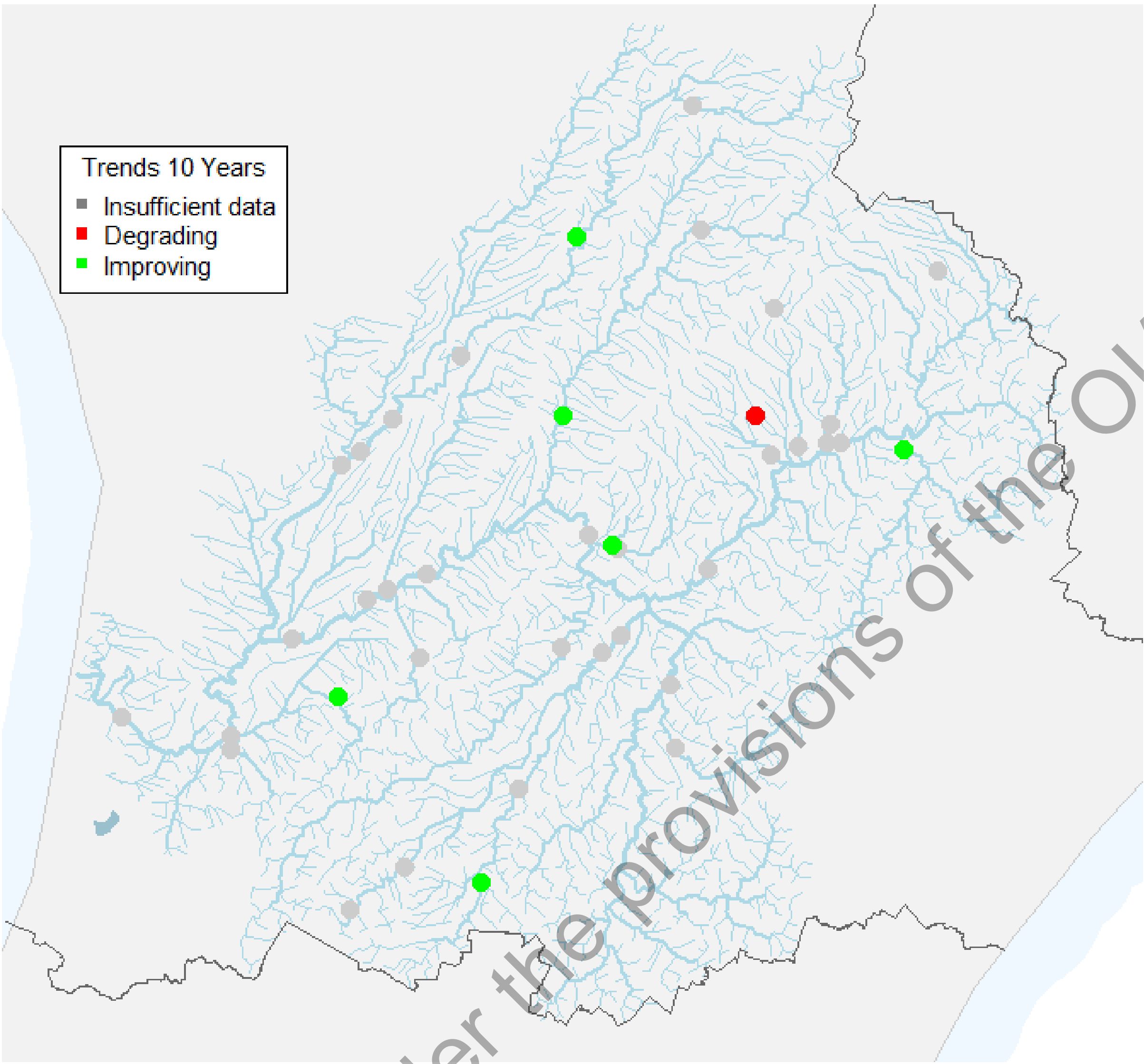

Figure 7. Map of the Manawatū River catchment DIN monitoring site locations with sites coloured by 10-year trend evaluations.

Trends in CHLA were evaluated for two time periods; the 5 and 10-year periods ending 2017. The results provided here are raw trends (i.e., not flow adjusted) because flows were not available for some sites. Fraser and Snelder (2018) showed that there are some differences between raw and flow adjusted trends for some sites, but that conclusions about general water quality changes (i.e., over all sites) were similar for raw and flow adjusted trends.

For the 5-year period, there were 29 sites with sufficient data for trend analysis of CHLA. Of these, no sites had degrading trends established at the 95% level of confidence and 4 sites had improving trends (Figure 8). All other (24) sites had insufficient data to establish trend direction at the 95% level of confidence. However, the aggregate trend for the whole FMU, as indicated by the proportion of improving trends PIT) indicated that 80% of sites were improving (standard error of 7.6%). This indicates that 20% of sites were degrading.

For the 10-year period, there were 23 sites with sufficient data for trend analysis of CHLA. Of these, eight sites had degrading trends established at the 95% level of confidence and no sites had improving trends (Figure 8). All other (15) sites had insufficient data to establish trend direction at the 95% level of confidence. The aggregate trend for the whole FMU, as indicated by the proportion of improving trends PIT) indicated that 17% of sites were improving (standard error of 7.3%). Note that this implies that 83% of sites were degrading over the 10-year period.

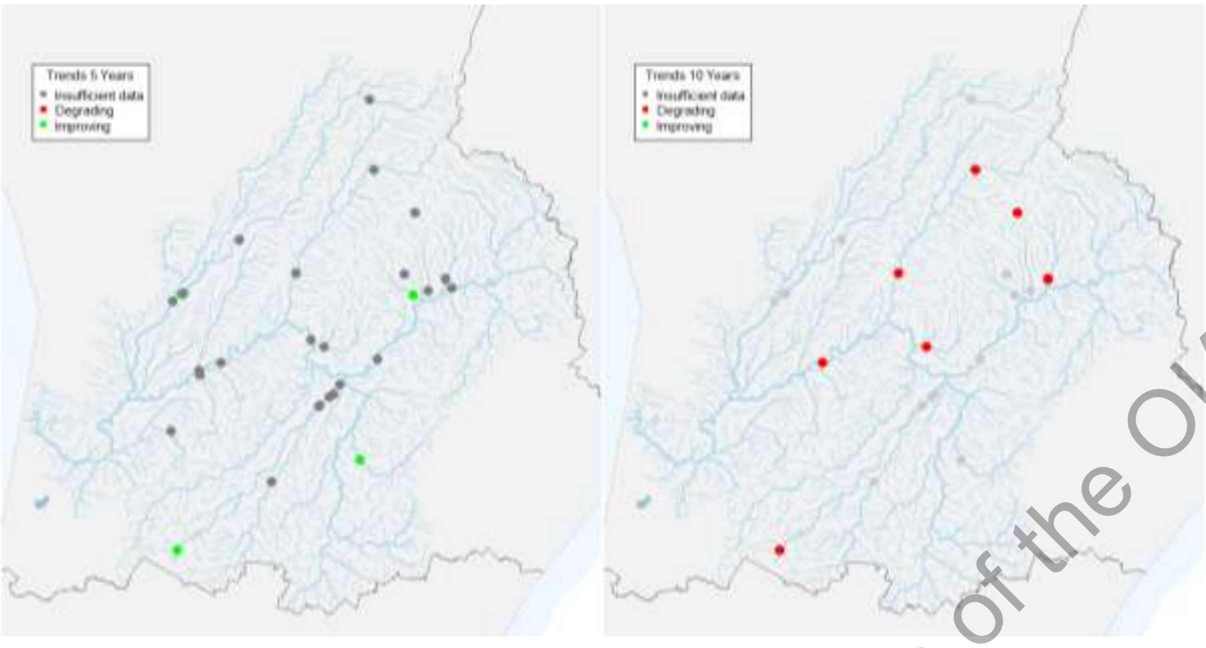

Figure 8. Map of the Manawatū River catchment periphyton monitoring site locations with sites coloured by the 5 (left) 10-year (right) trend evaluations.

Discussion

State

The 95% confidence intervals for the median and the 95th percentile NO_3-N concentrations were generally much smaller than the width of the A and B bands (Figure 1). This indicates that setting objectives based on NOF bands could allow degradation of water quality from current state up to the band threshold.

The state results must be examined in the context of the regional plan objectives. The analysis indicates that DIN targets are not being achieved at 35% of sites. These are sites at which improvement is required by the regional plan but judgements about whether the plan is succeeding would need to consider the timeframe for that improvement set out in the plan. Over time, the expectation would be that a greater proportion of sites within the FMU would achieve their targets. The rate at which this improvement should occur is not a scientific decision, it is a judgement that may be explicit in the plan or may need to be made as part of the assessment of maintain and improve.

The purpose of the DIN target is to achieve acceptable river trophic states, which are defined by objectives for periphyton biomass measured as CHLA. The state results for CHLA indicate that five of 18 the sites with a NOF A and B grade for CHLA failed to achieve the DIN target. This contrary outcome reflects the complicated nature of river ecosystems and the difficulty of setting criteria for DIN. Judgments about whether sites are maintaining or improving water quality when they are achieving their CHLA objectives but not achieving their DIN target is not a scientific decision. This judgment may be explicit in the plan or may need to be made as part of the assessment of maintain and improve.

The question of whether DIN targets are being achieved at individual sites is also complicated by the unavoidable uncertainty of the assessment. Of the 46 sites for which state was assessed, there were 12 for which the 95% confidence interval of their mean DIN concentration included the target (Figure 4). This means that the site gradings are equivocal at the 95% level of confidence for these sites. The site gradings could be unequivocally defined for some of these sites at lower levels of confidence, but a decision about the required level of confidence is needed. This is not a scientific decision because it reflects judgments about the burden of proof and management implication of getting the grading wrong.

The NO3N results highlight an important issue associated with aggregating data over many sites within an FMU. The assessment indicates that 79% of sites in the FMU are assigned to the A band for the NO3N attribute and the remainder are assigned to the B band. It is unclear how the performance of the FMU with respect to NO3N should be graded. This is further complicated by the fact that the assignments for individual sites are themselves uncertain. Various quantitative options are possible including basing the assessment on an average over all sites, a weighted average with the weighting attempting to adjust for bias in the site network or basing the overall grading on the worst case. The choice among these options is not a scientific decision because it reflects judgments about the burden of proof and management implication of getting the grading wrong.

Trends

The trend results for DIN indicated a majority of sites (i.e., pan-FMU) improved over the 10-year period (indicated by PIT statistic of 71% ± 11.6% at 95% confidence), but a majority of sites degraded over the 5-year period (indicated by PIT statistic of 11% ± 6.1% at 95% confidence). It is unclear how this reversal should be interpreted. A naïve interpretation could be that changes in the last 5-year period are undoing the benefits of changes made between 5 and 10 years ago. However, trends analysis indicates nothing about causes and we know from work by Scarsbrook *et al.* (2003) that these trends are likely influenced by climate variation. The trends therefore tell us that there have been changes in water quality but tell us nothing about whether these are due to management actions.

The trend results for CHLA indicated a majority of sites (i.e., pan-FMU) degraded over the 10-year period (indicated by PIT statistic of 80% ±14.9% at 95% confidence), but a majority of sites improved over the 5-year period (indicated by PIT statistic of 17% ± 14.3% at 95% confidence). The patterns in trends for CHLA were therefore the reverse of the patterns for DIN. Given that theoretically, DIN trends would drive trends in CHLA, it is unclear how this apparently conflicting pattern should be interpreted.

At the level of individual sites, a large proportion of trend assessments are equivocal about trend direction at the 95% level of confidence. The trend directions at individual sites could be unequivocally defined for some of these sites at lower levels of confidence, but a decision about the required level of confidence is needed. This is not a scientific decision because it reflects judgments about the burden of proof and management implication of getting the trend direction wrong.

Additional considerations

In order to keep this example simple, the analyses were based on current data pertaining to state of environment sites only. The analysis therefore reflects general condition and changes in conditions in the catchment at the current time. There are some additional issues that a comprehensive assessment of whether water quality has been maintained or improved would need to address.

First, some management actions address the upgrading of discharges from point sources such as sewage treatment discharges. The water quality downstream of these discharges is also monitored by Horizons and would need to be incorporated in the assessment.

Second, state of environment monitoring sites are generally biased to locations with poorer water quality and potentially a higher probability of degrading trends. Additional analyses such as spatial modelling can be used to produce a more representative assessment of current state and trends and this additional effort may be appropriate in some situations.

Third, water quality observations reflect the water passing a point in the catchment at a point in time, but may not represent the long term water quality that the catchment in its current configuration would achieve. Lags in the travel of contaminant through the drainage system (e.g., nitrate through ground water) may mean that future concentrations will be different to current even if nothing were to change in the catchment. The benefits of some management interventions may also be lagged (e.g., benefits of planting and soil conservation measures) and it may take many years before these are realised in the water quality observations. A comprehensive assessment would need to consider and address these issues to the extent that this is possible.

Implications and questions for proposals to change the NPSFM:

The analysis shows that current state can be determined at a site from water quality data. The evaluated state will often be more precise than a NOF band (i.e., the uncertainty associated with the estimate of state will be less than the width of a NOF band; Figure 1). This means the evaluated state can be used to set a more precise objective than a NOF band.

The analysis shows assessments of current state and trends based on monitoring data is always associated with uncertainty and may expose patterns that are inconsistent with general expectations. This uncertainty means that for some sites, we cannot be sure that an objective is being achieved or whether water quality improved or degraded. Therefore, any retrospective analysis of data that has the purpose of answering the question of whether water quality is being maintained and/or improved (e.g. when reporting on progress and reviewing plan efficacy) will be equivocal on multiple levels:

1. Establishing a state or trend at a single-site;
2. Establishing a state or trend at the FMU-scale – e.g. multiple sites may show a mix of improvement, decline, or indeterminate change.
3. Establishing a state or trend in water quality in general (i.e., across multiple variables). Note this example focused on one of many variable impacting on ecosystem health and other variables will likely show different, possibly opposing, patterns.

We note that assessment can be less equivocal about states and trends if we accept lower confidence requirements, and if we have more frequent monitoring/better data. However, there can never be complete certainty. In addition, we note that the required level of confidence is not a scientific decision because it reflects judgments about the burden of proof and management implication of getting assessments wrong.

We also note that the analysis of monitoring data described above do not give us information about their underlying causes states and trends. This is important in the context of reviewing plan efficacy. Establishing cause and effect in the context of environmental management is extremely difficult and it is likely that the effect of management actions will always be uncertain even with more intensive monitoring than what is currently the norm.

Specific questions

Is the STAG satisfied that setting freshwater objectives at or above the current state evaluated from monitoring data is technically feasible, and desirable, as opposed to defining state based on attribute bands?

We can deal with ambiguity with respect to whether water quality is being maintained and improved in two distinct ways – focus on reducing ambiguity (e.g., through more intensive monitoring, setting specific objectives capable of pass/fail analysis), or, focus on how to proceed despite ambiguity (e.g., acknowledging councils will have to make subjective judgements and providing them with guidance for this process). Both are valid, and the sub-group's approach has elements of both, but how far can/should we go with either approach?

Are there other opportunities to provide certainty through the NPS-FM and avoid subjective judgments (e.g. like requiring more specific objectives)?

Does uncertainty about both current state and changes as well as causes and effects mean that some subjective judgments are unavoidable?

Does the above mean that we should accept that it is not possible to define a 'bright line' test for whether water quality has been maintained?

What do different results at multiple sites mean at the catchment scale (e.g. NO3-N concentrations are spread across A and B bands)?

Is there an objective way to interpret this? To what extent does this require subjective judgment?

Is it useful for understanding what results mean at the catchment scale, despite any subjectivity?

What information is needed to enable councils and communities to answer the question of whether they have maintained water quality, as best they can? The paper summarising the outcome of the sub-group has some, but are these sufficient?

References

Fraser, C. and T. Snelder, 2018. State and Trends of River Water Quality in the Manawatū-Whanganui Region. LWP Ltd, Christchurch, New Zealand.

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

Scarsbrook, M.R., C.G. McBride, G.B. McBride, and G.G. Bryers, 2003. Effects of Climate Variability on Rivers: Consequences for Long Term Water Quality Analysis1. JAWRA Journal of the American Water Resources Association 39:1435–1447.