Freshwater Science and Technical Advisory Group:

16 April – priority paper compilation

Paper Author	Various	Classification	**Confidential**

Meeting date	16 April 2019	Agenda item (number)	

Contents:

Agenda Item	Title	Page
-	Agenda for 16 April	2
1	Action register	3
1	Draft minutes of meeting on 26 March	5
2	Nitrogen leaching - no paper	-
3	Rural package - no paper	-
4	Sediment	19
5	Wetlands	24
6	Nutrients	30
7	Ecosystem health metrics	57
8	Flows - no paper	-

Science and Technical Advisory Group Meeting

Agenda

Dates and Location: Tuesday 16 April 2019 9.30am-4.00pm, Terrace Centre Centre, 114 The Terrace, Wellington

STAG Members present: (TBC) Ian Hawes, Clive Howard-Williams, Jenny Webster-Brown, Ken Taylor, Bev Clarkson, Bryce Cooper, Jon Roygard, Russell Death (tentative), Adam Canning, Marc Schallenberg, Chris Daughney, Mike Joy, Ra Smith

Apologies: (TBC) Joanne Clapcott, Graham Sevicke-Jones

Items:

	9.00 am	*Coffee and tea*	(30 mins)
1.	9.30 am	Previous meeting minutes and actions arising, apologies, feedback from other advisory groups (Ken Taylor)	(15 mins)
2.	9.45 am	Nitrogen leaching (Claire Graham)	(15 mins)
3.	10.00 am	Rural package (TBC)	(15 mins)
4.	10.15 am	Sediment (Stephen Fragaszy)	(1 hour 15)
5.	11.30 am	Wetlands (Helli Ward)	(30 mins)
	12.00 pm	*Lunch*	(45 mins)
6.	12.45 pm	Nutrients (Jen Price)	(90 mins)
7.	2.15 pm	Ecosystem health metrics	(45 mins)
	3 pm	*Afternoon tea*	(10 mins)
	3.10 pm	Ecosystem health metrics continued	(20 mins)
8.	3.30 pm	Flows	(30 mins)
	4.00 pm	*Meeting close*	

FOR STAG CONSIDERATION ONLY NOT GOVERNMENT POLICY

Meeting date	Action	Who	Due date	Status
18-Oct-18	Officials to keep group up to date with climate policy developments	MfE	Ongoing	Incomplete
27-Feb-19	Commission research on extent and effects of superoxygenation in all ecosystems, and deoxygenation of lake hypolimnia.	Jen Price	? Long term	Incomplete
26-Mar-19	Conduct a case study of how proposed dissolved oxygen attribute would be implemented	Jen Price		Incomplete
26-Mar-19	Sub-group recommendations to be finalised	STAG sub-groups, MfE	16-Apr-19	Incomplete
26-Mar-19	Provide a worked example of a catchment to show how the different attributes relating to nutrients would fit together.	STAG sub-group, MfE	16-Apr-19	Incomplete
26-Mar-19	Provide advice on how uncertainty is taken into account in NOF attributes: 1. In the face of uncertainty how much of a margin are we giving to the environment, and 2. How confident are we that the number will provide the intended level of protection (may be qualitative).	STAG, MfE	?	Incomplete
26-Mar-19	Develop principles on uncertainty for attributes	STAG, MfE	? Longer term	Incomplete
26-Mar-19	Communicate uncertainty in attribute tables (may be qualitative).	STAG, MfE	When attribute tables are put forward	Incomplete
26-Mar-19	Incorporate suggestions on wording into sediment attributes	Stephen	16-Apr-19	Incomplete
26-Mar-19	Make clear that continuous monitoring of turbidity can be used to assess the suspended sediment attribute and is encouraged	Stephen	16-Apr-19	Incomplete
26-Mar-19	Email comments on exceptions to Stephen	STAG	16-Apr-19	Incomplete
26-Mar-19	Sediment discussion at next meeting	STAG	16-Apr-19	Incomplete
26-Mar-19	Case study: how do we manage a catchment when there are different classes present? For next meeting	Stephen/Jon	16-Apr-19	Incomplete
26-Mar-19	Provide more information to the group on a more sensitive macroinvertebrate indicator such as average score per metric	Carl		Incomplete
26-Mar-19	Investigate developing an attribute table with bottom line and bands for average score per metric, including what is the national state of this metric, how many waterways do not meet the bottom lines, implications for monitoring	?		Incomplete
26-Mar-19	Continue Ecosystem Health metric discussion by email and at next meeting	Carl/All		Incomplete
26-Mar-19	Discuss fish metrics further at next meeting	All	16-Apr-19	Incomplete

FOR STAG CONSIDERATION ONLY NOT GOVERNMENT POLICY

Meeting date	Action	Who	Due date	Status
26-Mar-19	Lake dissolved oxygen: Collate all the data available to assess current state and how many lakes would be below the bottom line, consider monitoring implications of new attributes	Lake DO sub-group working on		Incomplete
26-Mar-19	Examine what guidance would be needed for lake dissolved oxygen	Lake DO sub-group working on		Incomplete
26-Mar-19	Collate existing data and development of attribute tables for ecosystem metabolism metrics	MfE to commission		Incomplete

Science and Technical Advisory Group Meeting

Minutes - DRAFT

Dates and Location: Tuesday 26 February 2019 9.30am-4.00pm, Meeting room 1C (Ahumairangi), Environment House, 23 Kate Sheppard Place, Wellington

STAG Members present: Adam Canning, Bev Clarkson, Bryce Cooper, Chris Daughney, Clive Howard-Williams, Graham Sevicke-Jones, Ian Hawes, Jenny Webster-Brown, Joanne Clapcott, Jon Roygard, Mike Joy, Ra Smith, Russell Death (by Skype until 11am). **MfE staff**: Jen Price, Jo Burton, Stephen Fragaszy (morning), Isaac Bain, Nik Andic, Carl Howarth (Ecosystem Health session), Kirsten Forsyth (Flows session)

Apologies: Dan Hikuroa, Marc Schallenberg, Tanira Kingi, Mahinga-a-rangi Baker

1. Previous meeting minutes and actions arising, apologies, feedback from other advisory groups

Ken thanked Bryce for filling in as chair.

MfE staff gave a brief update feeding back the comments from the Regional Sector Water sub-group, and outlining the proceedings of the Freshwater Leaders Group (FLG).

Comments on the minutes from 26 February:

- STAG recommends further guidance is provided on the management and monitoring of dissolved oxygen (MfE staff are progressing this).
- MfE needs to provide a case study of how the proposed dissolved oxygen attribute would be implemented.
- Flows – It should be made clear that STAG suggested that changes should be made to the policies and objectives rather than incorporating a narrative attribute table.
- There was discussion about the management responses to different policy mechanisms. When talking about attributes, we need to consider the attribute development criteria which include linkages between attributes and the management actions. However, there are other policy options for incorporating things in the NPS that would not have this requirement. It can be useful to identify nationally mandated measurement requirements.

Outcome: The minutes from 26 February were approved.

Jo B outlined that the final decisions need to go to Ministers at the end of April. Everything we want to confirm needs to be sorted at the next meeting on 16 April. Sub-groups need to have their work done for the wider group to consider by the next meeting. The group should focus on the critical thresholds that we want to see in policy and not whether it should be an attribute or something else.

Actions:	For:
Conduct a case study of how proposed dissolved oxygen attribute would be implemented	MfE
Recommend guidance is developed on management and monitoring of dissolved oxygen	MfE
Update minutes from 26 February	MfE
Sub-group recommendations to be finalised	Sub-groups, MfE coordinators

2. Nutrients – brief report back on progress

MfE staff provided an update on the scoping of further investigation into Russell Death's proposed attribute tables for nitrogen and phosphorus. There was discussion around the technical details of the investigation.

Discussion points:

- It was clarified that the intent is for the most stringent nutrient attribute to apply.
- We need to be clear that managing nutrient objectives on their own is not enough to provide for ecosystem health.
- Managing nutrient values is something we want to strive for to manage ecosystem health. The effect on periphyton is not the only mechanism we need to consider. The nutrients will have an effect somewhere, in downstream receiving environments such as estuaries or the ocean.
- It was recommended that MfE analyses where the new attribute tables would take effect
- It was pointed out that there may be differences between data sets arising from different sampling methodologies. Another member has collated Macroinvertebrate Community Index (MCI) data as part of the sediment work, which examined the spatial variation in MCI. Land use or nutrients always account for over 2/3 of the variation in MCI. The remaining 1/3 of variation is influenced by temperature, substrate, and flow variability.
- It was noted that the proposed nutrient attributes only use MCI as one of many lines of evidence. All of the components give roughly the same numbers for nutrient attributes.
- We need to understand to what extent we over-or-underestimate the impacts.
- One member asked, what is the mechanism for a given nitrate concentration having a different effect in different classes, for example in Northland and Southland? It was noted that MCI reference state will vary between river classes. But we don't know the mechanisms causing these differences in MCI – it could be that one class is more impacted than the others.
- It was asked whether you could include a note accompanying an attribute table saying that councils should manage for the most sensitive ecosystem component?
- It was noted that the requirement to "maintain or improve" will also have an effect.
- Loss of riparian vegetation is a key factor influencing ecosystem health, this is something councils can manage.
- Summary from the chair: The group has some questions remaining around spatial variability, but we've received information that spatial variation in MCI is a smaller determinant for the variation. But we are looking at multiple lines of evidence. Suggested as a way forward – to proceed with a one-table approach, with an investigation of spatial variation to investigate whether it will be effective in achieving outcomes.
- One member pointed out that he has repeated the analysis in Russell Death's paper using quantile regression, this is a useful method for examining relationships when there is variability in responses. The quantile regression results broadly line up with Russell's numbers.
- It was recommended that a MfE provides a worked example of a catchment using all approaches. It would be good to provide three or so examples. It would be useful to look at soft-bottomed rivers where there is no periphyton.
- A further question was raised: What is the general approach for dealing with uncertainty? You can build precaution into the values in the table, or into the measurements. You don't need to do both. It seems that there isn't a generally agreed stance in the National Objectives Framework (NOF) about where uncertainty is built in. It would be good to have

an agreed approach. Action for MfE officials to provide advice on this and to advance work on principles to guide attribute development.
- It was recommended that the working group come up with an estimate of confidence in the numbers, e.g. for a given threshold, what is the mis-classification risk (ie, you think the ecosystem health objective will be met, but it won't). This can be expressed in qualitative terms if needed.

Outcome:

- It was agreed that STAG supports proceeding with nationally applicable attribute tables, with an investigation of spatial variation in the relationships between nutrients and ecosystem components. Another aim of the analysis will be to investigate where the attribute will be more constraining than the current periphyton attribute and accompanying note.
- The sub-group was asked to quantify uncertainty in the attribute table.
- STAG asked the sub-group to also provide a worked example of a catchment to show how the different attributes relating to nutrients would fit together.

Actions:	**For:**
Investigation of spatial variation in relation to where "unders and overs" might occur in relation to other existing nutrient attributes and MCI scores	Nutrient sub-group
Provide a worked example of a catchment to show how the different attributes relating to nutrients would fit together.	Nutrient sub-group
Provide advice on how uncertainty is taken into account in NOF attributes: 1. In the face of uncertainty how much of a margin are we giving to the environment, and 2. How confident are we that the number will provide the intended level of protection (may be qualitative).	MfE, STAG
Develop principles on uncertainty for attributes	STAG
Communicate uncertainty in attribute tables (may be qualitative).	Nutrient sub-group

3. Sediment

MfE staff asked STAG members for their feedback on the sediment attribute development work. This will be discussed again at the next meeting.

Questions discussed were:

a. Is the primary method on which bottom lines and bands are based – the community deviation method described in full in Appendix J – robust?

-STAG was broadly in agreement with the methods used.

Discussion points:

- 5 NTU represents good clarity, thresholds below this number may be within the margin of error of our measuring techniques.
- Foraging distance for fish will be impacted at the lower end of the turbidity measurements.
- Native fish forage at night and so are less impacted by turbidity than daytime foragers such as trout.

- A key point is that the proposed attribute states refer to medians and they do not represent the turbidity distribution at the site.
- Fish were the most sensitive species and so have been used to derive these values. • Measurements have been based on the NEMS methods.
- Members made recommendations for amendments to the attribute tables which are outlined in the Actions section below.
- When you look at water you can't tell the difference between 1, 5, or 7 NTUs. There is a perception issue, there is a risk this will be seen as false precision. But there are many measurements sitting behind these measurements.
- This highlights the need for and importance of continuous monitoring.

Discussion about the method:

- This is a bespoke method developed for this project.
- It's important that the method used to derive the reference state is clearly set out
- To weight the evidence, the team held a workshop with an expert from the USEPA and weighted the methods using set criteria. The research group has expressed uncertainty around where to set the thresholds, they chose to do that based on what has been done internationally.
- Why has the median been used? Because this can be linked back to the load from the catchment.
- When communicating this work it will be important to express why a new method is required for NZ.
- If you plotted the data distribution of turbidity observations, what would this look like for classes with very different medians? It is very variable between sites. The median averages this out. Any dose-response relationship will have this kind of variability.
- We should also communicate that a literature review has been done of impacts on sediment, to indicate a potential range of thresholds. International thresholds tend to be based on infrastructure management such as providing flushing flows.
- It was noted that there is international literature on acute effects but it's difficult to summarise because different measurement methods are used, and concentrations are quite high because they are based on 24-hour exposures. Therefore the group decided to focus on environmental ambient conditions.
- The bottom lines do not protect salmonids at particular sensitive life-stages. However, brown trout were incorporated into the analysis as an indicator. One of the principles for this work was to use native fish responses, but to use salmonids when insufficient data is available on native fish.

b. Can we provide for ecosystem health by including NPS-FM amendments with a deposited sediment attribute and a suspended sediment attribute using turbidity only?

-STAG was comfortable with progressing turbidity as the only attribute for suspended sediment. An attribute for water clarity is not required at this stage.

Discussion points:

- Clarity is more complex to measure than turbidity as it is influenced more by colour etc.
- It was suggested that examples could be given to show people what different turbidity levels look like.

c. Are the bottom lines set appropriately to provide for ecosystem health (keeping in mind the definition of the bottom line threshold)?

-STAG has recommended that it would be helpful to provide more description about whether the attributes do and don't protect, particularly in relation to salmonid spawning. STAG did not raise concerns about the bottom lines.

Discussion points:

- There was a discussion on the adequacy of the attribute for deposited sediment to provide for salmonid spawning. There is a proviso in the deposited sediment attribute table noting that all life stages of trout may not be provided for by this attribute tables.
- Fish & Game has developed a model of salmonid spawning reaches, we could examine the deposited sediment levels at salmonid spawning sites. Salmonid spawning will naturally not occur in all places.
- It was suggested that it would be helpful to include in the advice an analysis of where salmonid spawning occurs, and whether the attributes would provide for this.
- It was noted that native fish live in interstitial spaces, so providing for that may provide for trout spawning.
- It was pointed out that salmonids don't spawn everywhere, so an attribute for protecting trout spawning would not be relevant everywhere. The method predicts reference state based on the local environment.
- This method is based on community composition and has not been designed to protect individual species.
- This work uses presence-absence of fish because this is consistently quantified across the country. We now have the fish sampling protocols which will improve our data on fish abundance going forward.
- It was noted that the definition of C band includes "high likelihood of some impairment". It was recommended that these words require some consideration to make sure they are consistent with other attributes.
- In some classes, the attribute states are nearly exactly the same. Is there scope to reduce the number of classes and to have a narrative description of what each SSC class means? It was clarified that this will be included in an Appendix. MfE staff have discussed the implications of having many SSCs with regional council staff, and how these might be incorporated into FMUS which are defined differently in different regions.
- It was suggested that before we simplify things, we should ensure that our evidence is robust as possible. Is it worth holding off making a decision until we have better evidence?
- Another question raised was, how do we manage a catchment when there are different classes present? It would be helpful to have a case study on this. Guidance is required. STAG needs to understand how this will work to be able to provide advice.
- The existing periphyton attribute requirements are a good analogy of how councils deal with multiple objectives in one catchment. Councils are doing a good job of managing this. The bottom of the catchment e.g. the estuary, will be the constraining point.
- The policy needs to specify how management decisions will be made at the site level. E.g. will you apply the most conservative attribute?

d. Is the proposed classification system fit for purpose considering how ecological response information was incorporated?

- Some classes are nearly the same in terms of attribute states, it was asked whether it was possible to reduce the number of classes and to have a narrative description of what each SSC class means. Guidance is required on how to apply the different classifications in a Freshwater Management Unit.

e. Should we incorporate bands even though fewer lines of evidence support setting band thresholds according to the classification system?

-STAG recommends keeping the bands.

-Some of the classifications don't have much of a difference between the bottom line and the A band, we may not be able to practically measure the difference between attribute states given that most sampling is monthly. You wouldn't be able to eyeball the difference in clarity between the A band and bottom line in these cases and it will be difficult to communicate this to the public. However, these differences will have relevance for aquatic life.

f. Are the indicator definitions and monitoring requirements appropriate?

-STAG agreed that the indicator definitions and monitoring requirements were appropriate, and flagged that it's important to make clear that continuous monitoring can be used to measure the turbidity.

Discussion points:

- The attributes were developed based on single measurements at many sites, and do not include repeated measures of sediment over time.
- It was asked whether these attributes be applied to continuous monitoring? The response was: The monitoring requirement does not preclude continuous monitoring. There would be large implications of requiring continuous monitoring, so this is not being progressed as a requirement, just a recommendation.

g. Is the suspended sediment exceptions regime appropriate?

-This was not discussed due to time constraints and the chair asked the group to email any comments to Stephen after the meeting.

Ton Snelder gave a presentation on "Suspended sediment loads to ESVs – analytical framework" which had been requested by STAG, and there was some discussion on the details of this.

Actions:	For:
Incorporate suggestions: The wording describing the bands needs refining, it is not necessarily the median turbidity that is impacting the biota. The notation of the A, B, C, D, bands needs to be consistent with other attributes, using < etc. Provide more description of what the attributes do and don't protect, particularly in relation to salmonid spawning	Stephen
Make clear that continuous monitoring of turbidity can be used to assess the suspended sediment attribute and is encouraged	Stephen
Email comments on exceptions to Stephen	STAG
Sediment discussion at next meeting	All
Case study: how do we manage a catchment when there are different classes present? For next meeting	Stephen/Jon
Add Ton's presentation to portal	Jen

4. 11.30 am Wetlands

A discussion on wetlands was in the agenda, but this was deferred to an email discussion due to time constraints.

5. Flows – brief report back on progress

Kirsten updated the group on the policy work on flows.

MfE staff have previously presented a narrative attribute table to help councils set water quantity objectives. It was decided this was not the best way to proceed, but we still want to provide more guidance on how to set flows taking into account ecosystem health and human health for recreation.

MfE is progressing policy to direct councils to set objectives for surface water quantity taking into account habitat, flow variability, the habitat needs of aquatic life, and needs of connected surface water bodies. There will be guidance produced on various matters this policy raises.

Discussion points:

- The intent of this work is to replace the Proposed National Environmental Standard on Ecological Flows and Water Levels (2008). But the guidance going along with the NES will remain. The draft NES has a hydrological rule of thumb which defines the numeric part of flow setting. However, there is no guarantee that this approach will be protective in all cases. This is why we need more work on determining the ecological effect of existing flow regimes.
- One member recommended that more advice needs to be sought from NGOs and others.
- Flood protection works have an influence on flows and this should be considered.
- Flow is a critical component of many aspects of ecosystem health that we are considering. STAG members were disappointed that numerical values for flows are not being progressed.
- It was noted that the approach in the draft NES, if followed, would result in more water being allocated than is currently the case, and the environmental outcome would be worse.
- An approach is required that will be protective, but flexible enough to apply to different areas.
- One group member was of the opinion that the draft NES numbers could be provided as a minimum standard.
- It was recommended that the wording “recognise and provide for” is a stricter requirement and more appropriate in this case than “consider”
- It was suggested that numbers would be more reliable than a narrative.
- Another group member recognised that this is about strengthening the objective setting requirements, and wasn’t necessarily advocating for numerical values.

Outcome: STAG has indicated that more work on flow is critical. There will be more technical work that can be put into the NPS or an NES at a later date. This round of changes is setting the scene for this.

Key discussion points were:

- The group recognises that these changes are a first step and considers that more work on flow is critical. Numerical objectives are more reliable than policies and narrative objectives, and the group would like to see more work progressed on numerical objectives, providing adequate protection while accounting for differences in conditions around the country.
- The group recommended that stronger wording such as “recognise and provide for” is used
- Revised wording will be provided to the group for comment.

Actions:	For:
Kirsten has provided updated policy wording for STAG to comment on by email	Kirsten/all

6. Maintain or Improve

MfE staff summarised the issues and the recommendations of the sub-group and presented a draft flow-chart setting out the process of determining whether water quality has been maintained or improved.

Key outcomes from the discussion were:

- Members suggested that there should be a requirement to have no material reduction within a band. This option should be investigated further.
- STAG members will consider the option of process 2a in Nik's flow chart ("Have all freshwater objectives been achieved?"), to advance a statistical measure.
- Nik's flow chart step 2b ("Evaluate whether water quality has been maintained more broadly") is about integrating a more holistic view and evaluating a broader range of information. STAG will consider this option further.
- Nik will email the sub-group making specific proposals, listing alternative options, and asking the sub-group to comment.

Detailed discussion points:

- STAG has previously that defining current state within a band is not adequate.
- One member pointed out that the way current state is defined varies between attributes. For example, human health bands for risk of infection are very narrowly defined. They might not be meaningful for people. In that instance it might be acceptable to stay within Band B. But other attributes, such as the toxicity attributes, are broader, and maintaining current state based on bands is less helpful. It was noted that the definition of what is an acceptable infection risk is a decision made by politicians.
- One member suggested that the requirement should be that a decline within a band should not be allowed for some attributes.
- There are uncertainties associated with different measurement techniques that need to be taken into account when assessing data against objectives. It was suggested that this could be done using standard error, and it was pointed out that trend analysis can be used
- If the band is suitably narrow, if the confidence interval intersects the boundary between bands, this could be interpreted as failing the test for that band. However, some of the bands are not sufficiently narrow for this.
- It was noted that it is not the group's role to decide if bands confer an adequate level of protection, but to describe the level of protection that a particular objective will provide.
- One member was unsure if it's helpful to require that water quality at an individual site should be maintained or improved. Sites can change over time for good reasons. There needs to be flexibility allowed for regional councils to change monitoring sites. Maintain or improve should be defined at the FMU level; which would comprise several monitoring sites. You could define "overall" by calculating the proportion of sites that are meeting the objective.
- MfE staff clarified that defining "maintain or improve" at a site level would be a way of simplifying the process and removing ambiguity.
- The chair summarised that the key points from this discussion are (1) how you define whether water quality has been maintained and how the width of bands can help define this, and (2) an approach to defining maintain within an FMU with multiple sites.

- It was suggested that one option would be to require setting objectives at a site scale, and timeframes for achieving that. It then is simple to see if water quality has been maintained. The risk might be that if you were measuring within measurement error, the objective might be too restrictive.
- Members of the community have an idea about what maintaining water quality means. We need to explain our approach in a way that is meaningful for the community.
- One member noted that STAG has previously agreed that more bands would not be recommended, and that it is not acceptable to allow some sites within an FMU to decline while maintaining water quality "overall"
- It was pointed out that in some places, the community upstream has different values from downstream communities. FMUs were not based on communities, they are based on hydrological catchments in most cases. Communities will not always agree on the objectives to be met.
- The chair asked the group, can you define "maintain" in a statistical way?
- One option could be to provide a definition based on trends. However, one issue with this is that with monthly samples, you need a lot of data points to reliably demonstrate a trend.
- Biophysical measurements may not always be the best measures of change. Social metrics can be more responsive to changes being made in management.
- Attributes specify summary statistics and minimum requirements for data points.
- The discussion was summed up as: We can have a statistical measure of maintain based on a single water quality measure at a single site, but this doesn't help with defining what is happening across different variables. Aggregating sites and variables is not addressed by this. Defining maintain or improve across an FMU is more complex. Water quality state and trends can help, but you also need to measure inputs and social change.
- This discussion relates to the accounting requirements in the NPS-FM. Including communities in the monitoring would lead to greater engagement. The definition of freshwater accounting can be widened to better include communities in the process.
- Bryce – Aggregation can mask important detail. Having the raw information available for communities is important for guiding decisions. It's not our role to set out how the information is combined or aggregated, this is for communities and local government.
- The sub-group concluded that more information is required to assess whether water quality has been maintained or improved. Accounting should provide you with the information about any changes in sources and inputs that could lead you to expect a change in water quality.
- It was noted that we don't want a black box approach. Community members just want a simple answer about whether the water quality is going up or down, and what is happening to the bug numbers.
- If we chose to apply maintain or improve to attributes that vary less over time, such as MCI, would that be a solution? Others felt that MCI and QMCI are quite variable over time.
- It was noted that an objective is a desired outcome to guide plans and is not a compliance limit.
- MfE staff presented a draft flow chart setting out a process of determining whether water quality has been maintained or improved.
- It was noted that to achieve objectives, it is necessary to understand the actions that are required in the catchment. Different management levers have differing response rates. Achieving improvements in water quality requires a range of actions, and some are not included in the NPS-FM. In the past we have focussed on symptoms of ecosystem health decline, rather than the drivers. We need to consider how the NPS-FM fits in with the broader picture and other policy.

- The sub-group came up with two recommendations. One deals with statistical measures and the other deals with reporting against plans, e.g. The Selwyn-Waihora plan includes a requirement that nitrogen limits are not exceeded.
- There are two levels for defining maintain or improve – reporting against statistics, and reporting against the plans.

Actions:	**For:**
Nik will email the sub-group making specific proposals, listing alternative options, and asking the sub-group to comment. Proposals to be considered are: a statistical method for determining whether water quality is maintained; a "no material reduction" standard rather than current state, and a broader evaluation of information – not just water quality.	Nik, Maintain or Improve sub-group

7. Ecosystem Health

MfE staff gave the group an overview of different policy instruments being considered and it was agreed that the STAG would not be recommending particular instruments, but could make recommendations relating to the types of evidence available and management responses.

Copper and zinc

MfE staff provided an update about copper and zinc work as requested by one of the STAG members.

- Advice has been provided to Ministers.
- An attribute is not being progressed because regional councils do not have the means to control sources of copper and zinc.
- These sources require central government regulation and this will be progressed in the next 18-24 months.
- Copper and zinc attributes will be progressed in the next tranche of changes.
- MfE has scoped further work on acute and chronic attribute tables.

Macroinvertebrates

Discussion points:

- One group member recommended QMCI is used as it is a more sensitive indicator of effect. Well-known examples are acid mine drainage and wastewater treatment plant impacts. MCI may not indicate change to the same degree. The bottom line could be the QMCI equivalent of an MCI of 90.
- In the last round of submissions NIWA and NZFSS recommended that QMCI would be a better metric than MCI.
- It was suggested that if further macroinvertebrate metrics were progressed, guidance on sampling techniques would be required.
- It was noted that QMCI varies a lot more among seasons, this is why many regional councils use MCI in their plans. MCI is more useful for SOE monitoring, whereas QMCI is useful for addressing the effects of specific things.
- Including QMCI would require more monitoring and would require extra resources.
- It was pointed out that different methods are required for measuring biodiversity and effects of flows. MCI is not the right tool for these, and just sampling in riffles and runs is not sufficient for these applications. We may need to revisit how to monitor macroinvertebrates for ecosystem health. It was suggested that we need a wider shift in the way we use macroinvertebrates for monitoring. This would need to be done at a national level.

- One of the members recommended that both MCI and % EPT abundance (the percentage of individuals from Ephemeroptera, Plecoptera, Tricoptera) could be used, as this would avoid confusion between the two metrics. MCI is a measure of organic enrichment, whereas % EPT abundance is less specific. Bands would need to be defined.
- It was asked whether there was another way to bring the ecosystem health components into the NPS-FM, rather than introducing a large number of piecemeal indicators? We could introduce a more informative index that includes many indicators. The Ecosystem Health Framework is being progressed as a way to improve freshwater management.
- The "Average score per metric[1]" is another way of defining ecosystem health based on macroinvertebrates. A similar index has been developed using % EPT taxa rather than % EPT abundance[2].
- Further work would be required to do a case study incorporating % EPT abundance into this approach.
- There was discussion on the importance of lining up the bottom lines of different ecosystem health attributes. Group members recommended harmonising the different attributes. Approaches for this might involve matching up bottom lines, or using deviation from reference state.

Outcomes:

- STAG recommends retaining the Macroinvertebrate Community Index (MCI) in the NPS-FM, and that an attribute table with bands and a bottom line is developed.
- STAG recommends progressing more work on a more sensitive macroinvertebrate indicator. Different indicators were suggested, such as the published "Average score per metric" which incorporates three different macroinvertebrate indices. More work is required before an attribute table with bottom lines and bands can be developed, however it would be straightforward to ensure that the thresholds line up with MCI. These metrics are not directly connected to specific management levers but are useful and sensitive indicators of change. Councils could develop management plans to improve scores. There may be a requirement to conduct more monitoring to effectively measure such an indicator. More information and a case study needs to be provided to the group.
- There was discussion on how to set bottom lines and there was agreement that the bottom lines for different ecosystem health metrics should match up/be harmonised.
- Carl will provide further information to the group and further the discussion by email.

Actions:	**For:**
Develop attribute table with bands for MCI	? Adam?
Provide more information to the group on a more sensitive macroinvertebrate indicator such as average score per metric	Carl
Investigate developing an attribute table with bottom line and bands for average score per metric, including what is the national state of this metric, how many waterways do not meet the bottom lines, implications for monitoring	?
Continue discussion by email and at next meeting	All

[1] Collier KJ. 2008. Average score per metric: an alternative metric aggregation method for assessing wadeable stream health. New Zealand Journal of Marine and Freshwater Research 42: 367-378.
[2] Clapcott J, Wagenhoff A, Neale M, Storey R, Smith B, Death R, Harding J, Matthaei C, Quinn J, Collier K, Atalah J, Goodwin E, Rabel H, Mackman J, Young R 2017. Macroinvertebrate metrics for the National Policy Statement for Freshwater Management. Prepared for the Ministry for the Environment. Cawthron Report No. 3073. 139 p. plus appendices

Fish Index of Biotic Integrity (IBI)

Discussion points:

- The matters discussed by the sub-group were summarised.
- One issue is whether to use a regional or national model. There are regional differences, but it's not certain whether they are natural differences or caused by land use. What is an achievable A band in one region may not be achievable in another.
- Fish IBI is more holistic than other measures such as MCI because it takes into account downstream conditions, such as dams. The method is established and published.
- It was noted that trout are in the IBI – some regions have included trout as an "honorary native" species in the IBI based on the fact that trout are an indicator of good ecosystem health condition. Other exotic species such as carp are indicators of poor conditions. Trout and salmon habitat is protected under the RMA.
- Fish IBI is based on presence/absence and takes existing conditions into account.
- It would need to be standardised in a national model and could be an online tool.
- The model will be updated, with two versions, including and not including trout
- It was suggested that Fish IBI could be incorporated in the same way as MCI is currently in the NPS-FM.
- It is also possible to incorporate different scores for different threat classification levels.
- The Fish IBI incorporates six components and during model development, there was an analysis done to make sure none of the components are redundant.

Outcome:

STAG recommends progressing work on developing an attribute table with bottom lines and bands. Adam will develop this with Mike Joy (who originally developed the Fish IBI), investigating the implications of including trout in the index or not, as an "honorary native species". This is because trout, though not native, are sensitive indicators of land use and can provide useful information for guiding management.

Actions:	For:
Create proposed attribute tables with bottom lines and bands based on including and not including trout and report these back to the group	Adam
Discuss further at next meeting	All

Dissolved Oxygen in Lakes

A discussion paper was provided by Ian Hawes, Clive Howard-Williams and Marc Schallenberg proposing two attribute tables relating to biogeochemistry and habitat for aquatic species. The attributes take into account the need to harmonise with other attributes and account for natural variation. These attributes would provide a good fore-warning about declining lake conditions before they become severely degraded. There would be additional monitoring requirements for many councils to implement this attribute.

Discussion points:

- The long-term solution to address deoxygenation is to reduce the productivity in the surface waters of the lake.
- There may need to be exceptions for allow for naturally deoxygenating lakes. For example, in Lake Tikitapu, the hypolimnion deoxygenates naturally.

- Is this an average measure or something that cannot be exceeded? This is a one-off measurement at the end of the stratified period (i.e. in the last month before stratification breaks down)
- In polymictic lakes you would likely need to have continuous recording of dissolved oxygen.
- It was noted that the minima in the proposed lake dissolved oxygen attribute are lower than the river dissolved oxygen attribute. Most lake organisms are more tolerant of depleted dissolved oxygen than river organisms, and have more opportunity to seek better conditions elsewhere.
- It is the intent that both lake dissolved oxygen attributes would need to be met.
- It was noted that many regional council lake monitoring programmes are relatively new, and that there are many lakes in the "D" category of existing lake attributes.
- Some councils do helicopter sampling of surface waters, this would not be adequate for the proposed dissolved oxygen attribute.
- Management levers for lake oxygen are the same as those for TN, TP, and chlorophyll a.
- DO in lakes is a good metric for providing forewarning before conditions get too bad.
- This attribute doesn't require too much more work to develop. Work needs to be done on Question 4 – collating all the data available to see what the current state is.
- A group member estimated that there are measurements for 60 lakes out of 3800.
- It would be informative to examine trends approaching thresholds.
- This is also a climate change issue.
- Is there a way to direct councils to focus their efforts on lakes that are likely to be an issue? This also came up in relation to the dissolved oxygen attribute in rivers. You could incorporate a surveillance philosophy.
- Exceptions might be needed for peat lakes in Westland, geothermal lakes, and undisturbed lakes that are anoxic.

Outcome:

STAG was supportive of this approach.

Actions:	**For:**
Collate all the data available to assess current state and how many lakes would be below the bottom line, consider monitoring implications of new attributes	MfE to commission / Lake DO sub-group working on
Examine what guidance would be needed	

<u>Ecosystem Metabolism</u>

The proceedings of the sub-group were summarised.

Discussion points included:

- Ecosystem processes (as a component of ecosystem health) are not represented in the NOF at present.

- There has been previous work on this topic including Young et al. (2008)[3] and Cawthron has reviewed the applicability of these indicators in wadeable vs. non-wadeable rivers. The bands were suitable for non-wadeable rivers.
- Ecosystem metabolism approaches are not used in legislation overseas yet, this is likely to be related to the relatively new development of continuous monitoring of dissolved oxygen. There is a new free database where people are loading their dissolved oxygen data, and there are free online servers to calculate gross primary productivity (GPP) and ecosystem respiration (ER).
- STAG were supportive of this work and recommended that it should be progressed.
- Cotton strips would be a method that would be ready to be used immediately. This measures cellulose decomposition potential, and gives a measure of productivity. This is a relatively cheap test. There are internationally accepted methods and protocols.
- Are high or low decomposition rates good? The tricky part is that it is not a linear response variable. Usually slow is good and fast is bad because it is driven by nutrient status. But zero decomposition would also be bad.
- There have been national studies and this method is used by three regional councils. Cotton strips can be used for wadeable and non-wadeable streams, there are no bands for this yet but these can be estimated. This has not yet been tested.
- In general we lack metrics for large rivers. Ecosystem metabolism metrics can be used in large rivers and address this gap. These metrics need to be viewed in relation to the river continuum.
- GPP and/or ER would be appropriate for non-wadeable rivers.
- Further work would be needed to develop attributes for these metrics. This would be suitable for the second tranche of work which would be developed over a 18-24 month time frame.

Outcome:

STAG recommends that further work is carried out to develop attributes for ecosystem metabolism.

Actions:	**For:**
Collate existing data and development of attribute tables for ecosystem metabolism metrics	MfE to commission

Summary from chair

The sub-groups need to provide clear recommendations at the next meeting on 16 April. STAG will need a clear list of what is a priority for this round of changes and what will be progressed in the next round.

[3] Young RG, Matthaei CD, Townsend CR 2008. Organic matter breakdown and ecosystem metabolism: functional indicators for assessing river ecosystem health. Journal of the North American Benthological Society 27 (3): 605-625.

Sediment

The purpose of this brief is to provide updates on our policy recommendations to Ministers and the sediment impact testing work being undertaken. We also provide background materials to support our discussion of how proposed sediment attributes can be implemented within the FMU context. We welcome feedback on any issue raised and have specific guiding questions within the sections.

1. *Policy recommendations*

Suspended sediment

We are recommending progression of the suspended sediment attributes using the classification system provided in the NIWA/Cawthron report. At this time, we have not recommended specific bottom lines or bands, and we have not recommended specific requirements for the attribute monitoring or grading period. This is because we are still "road-testing" the proposed attributes with observed monitoring data and conducting impact-testing research.

In relation to the "road-testing", we will be focusing on results of continuous monitoring and also observation data from monitoring sites in reference or near-reference conditions per the NIWA/Cawthron definitions. In relation to impact-testing, currently we are testing whether the NIWA/Cawthron proposed bottom lines (community deviation method) are physically possible to meet across the country – whether land use and management changes can provide the suspended sediment load reduction required to meet the bottom lines.

If the research determines meeting bottom lines is impossible in some cases, we will examine whether that occurs only in specific classes, or across all classes. Depending on the results, we will evaluate whether and how the bottom lines and bands produced through another method, such as the extirpation method, alleviate the identified problems. Ultimately, if Ministers accept our recommendation to progress suspended sediment attributes, we will be advising on recommendations of specific bottom lines and bands as produced through the community deviation method or through another method.

Deposited sediment

We are recommending progression of deposited sediment monitoring plans and associated requirements for development of methods to address indicator scores below the NIWA/Cawthron-proposed bottom lines or declining indicator trends (hereafter monitoring plans). This will be comparable to the current treatment of macroinvertebrate community monitoring. We are recommending progressing monitoring plan requirements incorporating the classification system and indicator scores provided in the NIWA/Cawthron report.

Fundamentally, we consider that meeting the objective of the regulatory intervention – to improve management of deposited sediment to provide for ecosystem health outcomes – will be met better through the introduction of monitoring plans than through attributes. This is due to several underpinning considerations, which are discussed below.

Insufficient knowledge of management mechanisms

Our research shows that we have poor understanding of the land-to-river interactions that drive fine sediment deposition beyond the local scale. We do not understand how manageable (e.g. sediment load and riverbank stability) and unmanageable (e.g. slope and in most cases hydrograph) land- and river-scape features interact. Our research on deposited sediment management shows that suspended sediment load is a statistically significant but not important predictor of deposited sediment.

As a result, we are not able to assess the required interventions to meet particular quantitative indicator scores at the national, catchment, and, in most cases, sub-catchment scale. While we can model predicted deposited sediment areal coverage, we cannot effectively model how changes in land management and riverbank modification would affect those indicator scores.

Attributes require management at the FMU-level

Attributes require objective- and limit-setting to occur at the scale of the FMU. In contrast to suspended sediment and other existing attributes, deposited sediment indicator scores do not "integrate" the cumulative effects of land use and management at the catchment scale because they primarily reflect local hydrological and erosion (reach and run-scale) processes.

Development of limits and their litigation

Because of the above-mentioned deficiencies in our knowledge of how to manage deposited sediment, we consider that requiring all councils to set FMU-wide resource use limits would be counter-productive to the ultimate objective of improving management of deposited sediment to provide for ecosystem health outcomes. The limit-setting process is contentious and therefore litigious. If central government were to require councils to set limits in the absence of adequate knowledge about the relationship between regulated actions and the intended outcome, it would introduce a significant barrier to action in addition to being a major financial burden on councils. Aside from additional legal burden, we would also effectively be requiring councils to conduct research on the links between land/river management and in-stream deposited sediment outcomes.

Deposited sediment monitoring plans as a pathway to desired outcomes

The introduction of monitoring plans will further develop the information base necessary to improve understanding of deposited sediment management drivers. The structure of the proposal will also address the fundamental policy gap we identified as a key source of problems in resource management regarding in-stream sediment: that councils do not set region-wide in-stream sediment thresholds to maintain ecosystem health. Monitoring plans will provide this information in a spatially explicit way and leave to local government the methods to use and their scale of application when monitoring shows that indicators are below specific thresholds or trends decline.

<u>Guiding question:</u>

Several members of STAG have previously indicated a preference for attributes to monitoring plan requirements in relation to macroinvertebrate communities. Do you have the same preference regarding deposited sediment, and if so, can you please describe why you have this preference? Also, can you please provide suggested changes of policy wording that may allay your concerns regarding monitoring plans and their effectiveness in achieving their aim?

2. *Impact testing*

We have concluded one component of our impact-testing programme: modelling to assess the annual median suspended sediment load reduction required to meet the proposed suspended sediment attribute bottom lines produced by the community deviation method. The primary results are reported in Figure 1 below at the catchment scale, and we have also received the information at the river reach scale. Exceptions to the attribute have been incorporated in relation to glacial flour only.

This information will be used as the basis for erosion/economic modelling to assess the viability and costs/benefits of interventions to meet the proposed bottom lines. The economic modelling will provide results at the catchment level (and higher aggregation), and the erosion modelling will be produced to attain the catchment level reductions and show results at the reach-level as well. That will provide us additional information to assess how achievement of bottom lines at the catchment scale is reflected at the river reach scale (overs and unders within a catchment).

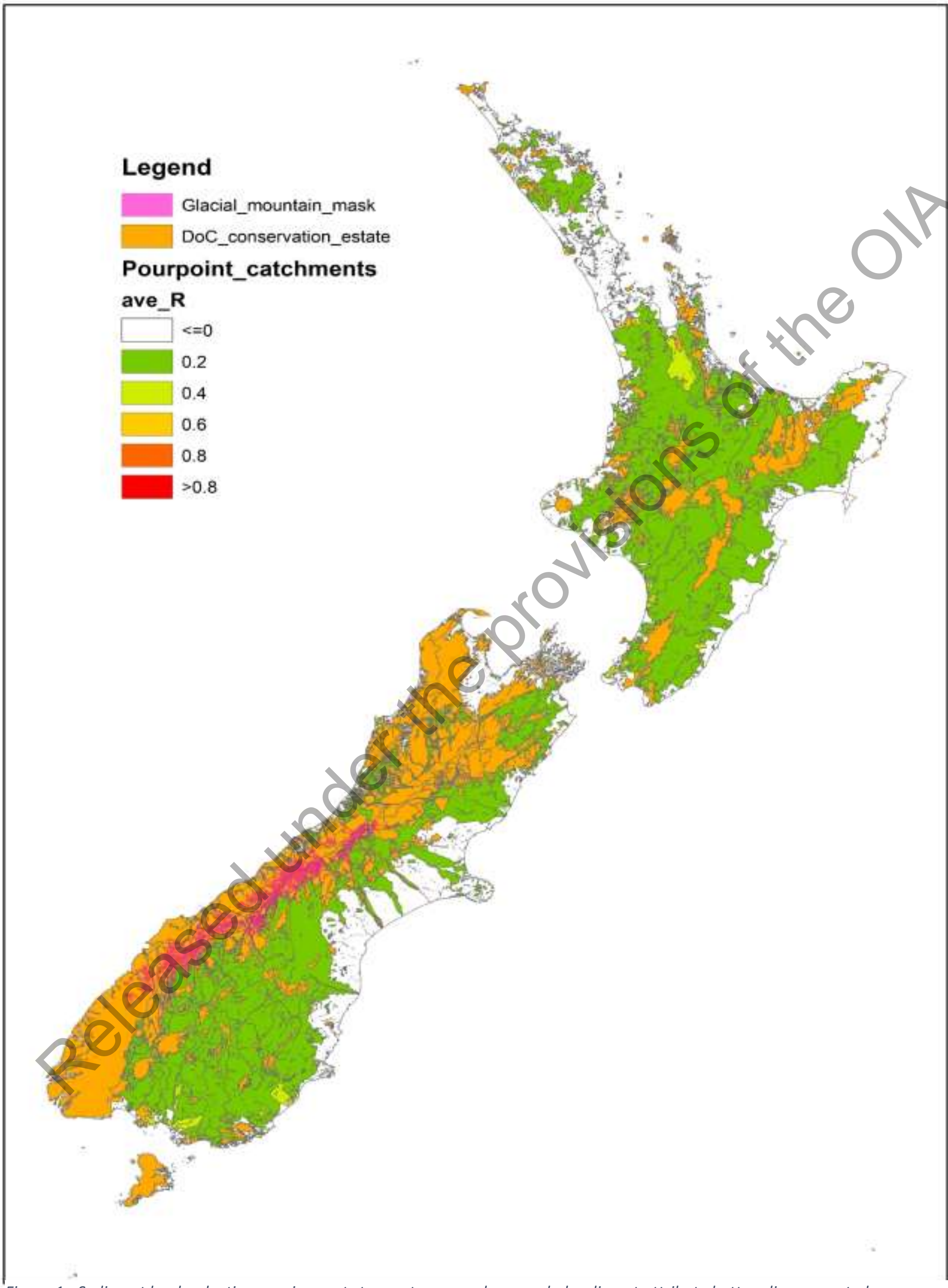


Figure 1 - Sediment load reduction requirements to meet proposed suspended sediment attribute bottom lines reported as "average R" which is the proportional annual load reduction per catchment.

3. *Sediment attribute implementation in the FMU context*

We would like to workshop with you how attribute implementation could occur within the FMU context. To aid the conversation, we will use a mapping application that shows a region, its FMUs, the proposed suspended and deposited sediment classification systems, existing SOE (turbidity) and deposited sediment monitoring sites, and long-term medians using available data. Through the conversation we would like to address the following questions:

A. Noting that attribute "grading" occurs at representative monitoring sites, is the monitoring network in the region and in FMUs adequate to support attribute and monitoring plan grading via the following three potential approaches:
    1. Blanket attribute grading across the FMU and sediment attribute classes (turbidity objectives in all classes must be in the B band);
    2. Attribute grading via classes (all areas in Class 1 must be in the B band, in class 2 in the A band, and in class 3 above the C/D band threshold).
    3. Attribute grading via percent of reaches achieving particular bands (20% of reaches in A band, 40% in B band, 40% in C band).

B. If not, where would additional monitoring be needed? This can represent coverage within specific FMUs, within specific sediment state classifications, or other characteristics.

C. If one monitoring site had continuous turbidity monitoring, would that change any of your comments above, and/or would it change how you would want attribute implementation characterised.

Wetland Water Levels – potential project

Problem

Altered hydrological regimes are causing the loss and degradation of wetlands in New Zealand. Evidence is required to provide guidance on the natural range of water levels for different wetland types, and what degree of variation relates to ecological risk including a decline in Ecosystem Health.

Background

The draft NES Ecological Flows and Water Levels (2008) recommendations for '*no change in water levels, beyond the water level variation that has already been provided for by existing resource consents*'.

This provision is problematic as it is open to subjective interpretation because there is inadequate definition of what is meant by water level variation, and many consents do not refer to specific levels which reduces the ability to apply the interim limit in practice.

The draft NES also proposed a risk assessment table that refers to the water level changes that reflect potential risk to wetlands[4]. This table is considered too broad to apply across all wetland types and a 0.2m variation may possibly be detrimental in some cases.

Project description

We are considering a joint technical project with DOC to describe the natural range of water level fluctuations of different wetland types, and determine the ecological risk associated with variations from annual and seasonal median water levels.

This project is in a scoping phase. We would like to present the preferred option to the STAG for advice on the level of technical detail experts are comfortable with to provide evidence to support policy.

The general approach would consist of:

- Collect and collate wetland level data (partly complete from an existing project)
- An assessment of water level data previously collated and any new data collected
- A literature review of ecological risk associated with water level changes for different wetland types
- An expert panel is convened to provide input on the ecological risk associated with variations from annual and seasonal median water levels.
- Recommendations (potentially using a matrix approach) from the findings of the above analysis of risk to different wetland types to water level change and an indication of how much certainty we have in the recommendations.

Note: The guidance from this technical project is intended to support councils implement proposed freshwater policy (eg, NPSFM and potential NES).

Questions

1- Do you agree with the general approach

[4] Low < 0.2m change in median water level; and patterns of water level seasonality (summer vs winter levels) remain unchanged from the natural state (summer relative to winter)
Med > 0.2m and < 0.3 m change…
High > 0.3 m change …

2- Within the various options over the page is there anything missing, anything you think is unnecessary, or anything you would mix and match for each level of effort?
3- Reflecting on the preferred option over the page what level of detail would you be comfortable with for providing guidance on wetland water level variation and the associated ecological risk?

Effort	Description	Timeframe and cost estimate	Comment
Low	1. Collect hydrology data on wetlands, grouped by wetland type (i.e. bog, swamp). Collated data already exists for ~30 wetlands across NZ as part of previous study. [Additional data from 10-20 wetlands may be required (may only require one transducer per wetland), which may extend project duration]. 2. Correlate the hydrology type with wetland classification, either via technical reports or expert classifications of the site (i.e. Transducer at X location is a Restiad Bog with 5 m of peat). 3. Analyse the 'grouped' wetland hydrological data to determine typical annual water level ranges, maximum/minimums observed. 4. Rank the wetland groups based on transition driven by changes in hydrology (i.e. bog – fen -swamp -marsh) and the degree of hydrological intactness of the wetland sites/groups. 5. Identify where the water level (WL) 'tipping point' occurs where water level fluctuations begin to move one wetland type into another (i.e. fen to swamp, bog to fen), or corresponds with an ecological shift. 6. Correlate tipping point with agreeance by a nominated panel of experts 7. Write a technical report documenting the WL tipping point, and include a literature review on national studies. 8. Peer review of report	Timeframe: 7-10 months, including peer reviews by experts assumed to take up to 3 months Cost: depending on extent of peer review and number of sites	PROS: 1. Can be undertaken immediately 2. Can potentially incorporate some numbers on water levels, and water level variation, that corresponds to ecological risk for NES updates by end of 2019 3. Lower cost 4. Method focusses purely on WL data, in a way, ignores SW/GW/RW inputs to simplify assessment CONS: 5. Difficult technical assessment if only focussing on hydrology (would be improved with an eco-hydrological assessment of soils + vegetation) 6. If fast tracked, may not identify all available data (i.e. international data) 7. Assessment may be for limited wetland types if fast tracked 8. Extent of assessment will need to be streamlined (i.e. focussing primarily on water levels)

- TBC: **Moderate and high effort**, assume to extend assessment including soil + vegetation correlations, wider literature review (international).
- This project is expected to align with the Landcare Research work on wetland delineation (hydric soils).

Wetland Drainage Setback – potential project

Problem

Altered hydrological regimes are causing the loss and degradation of wetlands in New Zealand. Drains within and around wetlands lower water tables. Evidence is required to provide advice on appropriate drainage setback distances from the edges of wetlands to avoid altering their natural hydrological regime.

Of those councils that stipulate drainage setback from wetlands distances vary from 25m to 200 m.

Project description

We are considering a joint technical project with DOC to provide evidence on the impact of drains on the water levels of wetlands and to recommend minimum setback distances for drains taking into account the water regime of different wetland types and different hydro-geological settings

This project is in a scoping phase and there are various options we could pursue over several stages. We would like to present the options of this and the wetlands water levels project to the STAG for advice on the level of technical detail experts are comfortable with to provide evidence to support policy.

The general approach would consist of:

- Collect and collate hydrological data associated with wetland drains (nb: several hydrological research projects have been established in New Zealand)
- Analyse data of effects of drainage in and around wetlands
- A national and international literature review of drainage impacts on wetlands
- Modelling water table draw down versus setback distance for different wetlands
- Field investigations to calibrate model
- An expert panel discussion to recommend minimum setback distances
- Recommendations of appropriate setbacks for drainage presented in a technical report

Some initial options are detailed over the page which reflect differing levels of effort and cost.

Note: The guidance from this technical project is intended to support councils implement proposed freshwater policy (eg, NPSFM, potential NES).

Questions

4- Do you agree with the general approach
5- Within the various options over the page is there anything missing, anything you think is unnecessary, or anything you would mix and match for each level of effort?
6- Reflecting on the options over the page what level of detail would you be comfortable with for providing guidance on setting appropriate setback distances from wetlands for drainage?

Table 1. Wetland Setback Options (for land drainage/ditches)

Effort	Description	Timeframe and cost estimate	Comment
Low	1. Literature review – compile national and international literature on wetland setback distances 2. Data collection-utilise suitable existing national wetland hydrological data available at the time of the study (est 5-7 sites) 3. Data analysis – review drawdown from drains and estimate the extent of the impact zone which results in degradation/transition 4. Compile a technical report recommending setback distances for the assessed wetland types, nominal recommendations for 'un-assessed' wetlands (i.e. gumland) 5. Have technical report and key tables peer reviewed	Timeframe: 6-7 months, including peer reviews by experts assumed to take up to 3 months Cost: low, depending on extent of peer review.	PROS: 2. Can be undertaken immediately 3. Can potentially incorporate some setback numbers in NES updates by end of 2019 4. Lower cost CONS: 5. If fast tracked, may not identify all available data 6. Assessment will be for limited wetland types. 7. Numbers may come under scrutiny as limited sites only looking at hydrology + literature (excluding vegetation + soils)
Med	1. Literature review – as above, collate data on wetland soils for modelling 2. Data collection-extensive data collection, contacting district/regional councils, CRI's (i.e. Landcare) and private organisations to compile suitable information (est 10-20 sites) 3. Data analysis – as above, but for greater sites and likely more wetland varieties 4. Model build – Create a simplified SeepW model representing a farm drain, calibrate the model to some observed data to test performance 5. Model scenarios – Run scenarios of a range of wetland types with different soil parameters to populate data gaps. 6. Compile results – From observed data and modelled 7. Compile technical report (as above) 8. Technical report and modelling review (model review to occur once model has been built, prior to scenarios being run).	Timeframe: 12-14 months, depending on discussions and data collection with councils, peer reviews etc. Cost: moderate, which will depend on the number of sites to be assessed and the extent of model scenarios.	PROS: 1. Thorough review of available national data 2. Greater confidence in setback distances, varied by wetland type and supplemented by robust peer reviewed modelling outputs 3. Greater representation of a variety of wetlands 4. Stronger confidence in the numbers that would be represented in national policy, less chance of scrutiny CONS: 5. Slower to complete, could be delayed by reviews and data procurement 6. Greater cost, which could increase due to the number of parties involved and timeframe 7. The numbers, while supported by modelling and greater datasets, do not include Soil + Vegetation assessments

Effort	Description	Timeframe and cost estimate	Comment
High	1. Literature review – as above, extending data collection for modelling + hydric soils following Landcare Research 2018 work on wetland delineation 2. Data collection – as above 3. Data analysis – as above 4. Model build – as above 5. Model scenarios – as above 6. Field Investigations – Conduct a number of field visits and soil/veg samples across a number of wetland types to validate the model scenarios (minimum 4). 7. Compile results – From observed hydrology data, modelled information and also soil + vegetation samples to link to degradation from drainage (possibly run some principal component analysis for statistical trends) 8. Compile technical report – as above 9. Technical report and modelling review – as above.	Timeframe: 12-18 months, depending on discussions and data collection with councils, people involved (i.e. collaborative efforts) peer reviews etc. Cost: high, which will depend on the number of sites to be assessed and the extent of model scenarios.	PROS: 1. Extensive assessment of available data (nationally) and globally (literature) 2. Thorough data review and compilation of modelling, soil, vegetation and hydrology information 3. External reviewed and strong confidence in numbers to brought into policy CONS: 4. Could take time to secure funding, and a longer time to complete the work 5. Cost (most expensive option), which could creep unless a robust scope is developed and held to 6. Delays (due to organising field work, identifying sites, peer reviews etc)

Nutrients

Context

STAG and Freshwater Leaders Group (FLG) have recommended two improvements to the way nutrients are managed in the Freshwater NPS to ensure that ecosystem health can be adequately maintained or improved. For the following recommendations, where there is more than one attribute for the same metric, STAG recommended that the more stringent one should apply.

Recommendation 1: Guidance tables for nitrate-nitrogen and dissolved reactive phosphorus – for different river classes

In addition to the current process set out in the Freshwater NPS, STAG members recommended providing as guidance an optional default set of tables for DIN and DRP in different types of river. Councils could choose to use the default tables or derive their own DIN and DRP objectives. The implication of this option would be to reduce the burden on councils to derive their own criteria for nutrients.

STAG and officials will work together to investigate and progress this recommendation. Work so far has focussed on the second recommendation as described below.

Recommendation 2: Attribute tables for nitrate-nitrogen and dissolved reactive phosphorus – applicable nationally

STAG also considered a proposal from Russell Death for new nitrate-N and DRP attributes for inclusion in the Freshwater NPS. These attributes differ from the default tables described in Recommendation 1 in that they have been derived through relationships with many ecosystem components, rather than just periphyton, and one table would apply nationally, rather than having different tables for different river types.

The group recommended that attribute tables addressing the effects of nutrients on a broader range of ecosystem components should be incorporated into the Freshwater NPS. STAG agreed that the methodology used to derive the suggested new attributes was robust, but there were some outstanding questions:

1. how to weight multiple lines of evidence
2. how and where to set bands in relation to ecological responses
3. whether attribute tables should vary spatially
4. whether total nitrogen (TN) and total phosphorus (TP) attributes for ecosystem health are more appropriate than nitrate-N and DRP

Below is a document containing analysis to inform further discussion on nutrients.

Questions to guide discussions

1. Does the evidence indicate that single, nationally applicable nutrient tables for N and P are tenable?
2. Does the evidence indicate that spatially varying nutrient tables for N and P are tenable?
3. What are the appropriate metrics? (DIN, DRP, TN, TP)
4. Based on the answers to the questions above, what are the limitations or caveats that need to be communicated?

Investigations of relationships between DIN and DRP and Ecological Health to assist STAG

Introduction

Death *et al.* (2018) proposed national thresholds for dissolved inorganic nitrogen (DIN) and dissolved reactive phosphorus (DRP) for the protection of ecosystem health. The thresholds are based on fitting regression models to site-median values of DIN and DRP concentrations and a range of measures of ecosystem health that included site-median MCI and IBI scores. Relevant thresholds for DIN and DRP were derived by first nominating a graduated range of target values for the measures of ecosystem health from A (very good state) to D (poor state). The corresponding DIN and DRP thresholds were obtained from the regression models by finding the concentrations that were associated with the nominated ecosystem health targets. A weight of evidence approach was used to summarise the multiple DIN and DRP thresholds to a single set of thresholds (Table 1).

Table 1. Proposed attribute states for nitrate and dissolved reactive phosphorus.

Attribute state	Annual Median (mg/l) (Nitrate-N)	Annual Median (mg/l) (DRP)
A	≤ 0.10	≤ 0.006
B	>0.10 and ≤0.46	>0.006 and ≤0.019
C	>0.46 and ≤0.89	>0.019 and ≤0.038
D	>0.89	>0.038

The STAG expressed two reservations with the approach. First, Death *et al.*'s (2018) analysis was undertaken at the national scale and therefore differences in environmental conditions that may lead to differences in the relationships of DIN and DRP with ecosystem health were not taken into account. Two types of differences may be expected. Ecological health measures may be expected to vary naturally (i.e., in addition to the response to DIN and DRP) across environments of different type due to factors such as local habitat, flow regimes and light and temperature. In addition, relationships of DIN and DRP with ecosystem health can be expected to vary because they are mediated by naturally varying environmental factors (e.g., local habitat, flow regimes and light and temperature). The second reservation expressed by STAG concerned the correlative nature of the relationship of DIN and DRP with ecosystem health. DIN and DRP concentrations generally vary longitudinally within a catchment in response to the intensity of land use upstream. However, many other variables also vary longitudinally (e.g., local habitat, flow regimes and light and temperature). There are many correlative relationships between ecological health variables and environmental variables, and there are likely to be multiple causal variables among the environmental variables. Death *et al.*'s (2018) analysis provide no evidence that DIN and DRP are the sole causal variables and it is reasonable to expect other environmental variables are among the causative agents. If variables other than DIN and DRP are among the causative agents, then actions to manage these nutrients will not produce the desired change in ecological health.

Currently STAG is considering the proposal that Death *et al.*'s (2018) national thresholds for DIN and DRP would have effect in cases where they were lower than concentrations that were deemed to apply under the existing national objectives framework (NOF) requirements. The NOF requires that regional councils set DIN and DRP criteria for the management of periphyton or, if a stream does not support conspicuous periphyton, concentrations of nitrate-nitrogen and ammoniacal-nitrogen must be set to levels to avoid toxic effects. The NOF requirements are referred to hereafter as the NOF DIN and DRP criteria. STAG also requested an analysis of the spatial impact of Death *et al.'s* (2018) DIN and DRP thresholds including where they would apply and by how much current concentrations of DIN and DRP would need to be lowered to comply.

STAG requested additional work to assist in their decision making and this paper describes analyses undertaken in response to this request. Three main sets of analyses are included:

1. Assessment of variation in the relationship of DIN and DRP with ecosystem health measures within river classes.
2. Assessment of the 'uniqueness' of the ecosystem health DIN/DRP relationships. This analysis cannot prove or disprove that DIN/DRP is the causative agent but can quantify the extent to which these relations may be overestimated if other environmental variables are not considered.
3. A national analysis comparing the impact of the NOF DIN and DRP criteria with the proposed thresholds of Death *et al.* (2018) based on estimated concentrations of DIN and DRP in all rivers nationally.

MCI scores have been used in analyses 1 and 2 as proxies for ecosystem health. It is acknowledged that the thresholds of Death *et al.* (2018) are based on numerous biological response variables including MCI. The analyses are therefore demonstrative rather than comprehensive.

Data

Ecosystem health and nutrient concentrations

An alternative dataset to that used by Death *et al.* (2018) was used in this analysis. Data describing observed MCI scores and concentrations of nitrate-nitrogen (NO3N) and DRP were obtained from a national database of regional council monitoring data and NIWA's NRWQN network. The dataset comprised NO3N rather than DIN, and because NO3N comprised the majority of DIN, the analyses that follow have been based on NO3N. For sites at which both MCI and the nutrients were monitored, the median of all observations in the period 2013 to 2017 were calculated. The median values were used in all subsequent analyses to represent ecosystem health (MCI) and nutrient concentrations (NO3N/DRP) at each site. There was a total of 450 sites with MCI and nutrient data that were well distributed nationally (Figure 2).

Some members of the STAG were concerned that the national dataset comprised a mixture of collection agencies and that differences in sample collection, processing and MCI score calculations may confound the analysis. To help to address this issue, one of the analyses (Variance partitioning) was performed on a smaller dataset comprising only 47 NRWQN sites at which invertebrates are collected. The NRWQN dataset is collected and processed by a single agency (NIWA) and is also national in extent.

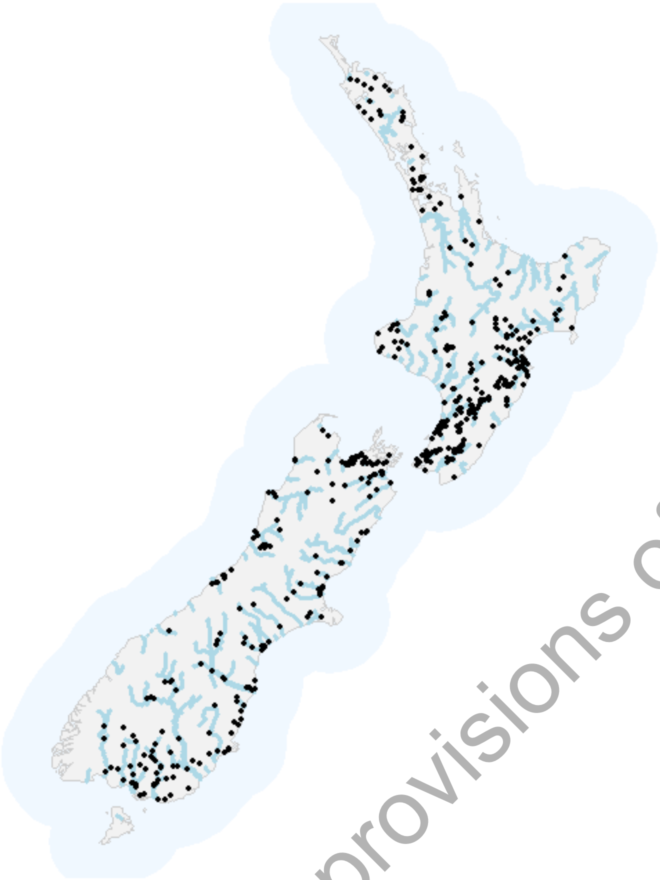


Figure 2. Location of regional council and NRWQN sites with MCI and nutrient data.

Classification of sites

The sites were classified using two national river classification systems, the River Environment Classification (REC; Snelder and Biggs, 2002) and the Freshwater Environments of New Zealand classification (FWENZ; Leathwick *et al.*, 2011). These classifications are associated with a national digital river network that comprises 560,000 segments (defined by upstream and downstream confluences) with a mean length of ~700m. The classifications are contained within a Geographic Information System (GIS).

The REC is a general classification that provides resource managers with a multi-level hierarchical classification of rivers, and an ecological basis for catchment management (Pyle et al. 2001). The first two levels of the REC are referred to collectively as Source-of-flow classes and define river classes based on differences in catchment climate and topography. REC Source-of-flow classes have been shown to broadly discriminate water quality, invertebrate communities, hydrology and river morphology. However, catchment land use is associated with Source-of-flow classes and STAG expressed concern that this association may confound the relationship of NO3N and DRP with ecosystem health.

The FWENZ is an alternative classification that was developed for biological management and monitoring/reporting purposes. The classification maximises the discrimination of natural patterns in fish and invertebrate communities in New Zealand's rivers (Leathwick *et al.*, 2011). The definition of FWENZ classes is therefore minimally influenced by catchment land use.

Additional environmental explanatory variables for sites

Additional environmental descriptors for the monitoring sites were obtained from the freshwater environments of New Zealand (FENZ) database. Each segment is associated with many descriptors that were derived by intersecting the network with other spatial layers as described by Wild *et al.* (2005).

Two groups of environmental explanatory variables in the FENZ database were chosen to represent the character of the segment and the upstream catchment of each monitoring site. These additional explanatory variables were derived from spatial layers including a terrain model, mapped climate data, and geological maps. The variables were chosen based on previous analyses that have demonstrated their association with ecological characteristics (e.g., Leathwick *et al.*, 2011). The additional explanatory variables included catchment climate (usAvTWarm and usRainDays10), catchment topography (usAveSlope and usLake) and the character of the catchment surface geology (usHard and usPhos). Four variables were chosen to represent the characteristics of the segment including its position in the network (ORDER) its elevation (segAveElev), slope (segSlope) and distance to the coast (dsDistToSea).

Table 2. Additional environmental explanatory variables (see Wild et al.*, 2005 for details)*

Type of variable	Variable name	Description (units)
Catchment variables	usAvTWarm	Mean January air temperature (°C x 10)
	usRainDays10	Catchment rain days (greater than 10mm/month) (days/month)
	usAveSlope	Average slope of catchment calculated from 30m DEM grid (m/m)
	usLake	Lake index (dimensionless)
	usHard	Catchment average of hardness (induration) of surface geology (ordinal)
	usPhos	Catchment average of phosphorous in surface geology (ordinal)
Segment variables	ORDER	Stream order (ordinal)
	segAveElev	Average segment elevation (m. asl)
	dsDistToSea	Mean January air temperature (°C x 10)
	segSlope	Average segment slope (m/m)

Estimates of current concentrations

Model predictions of current concentrations of NO3N, TN and DRP were obtained from Larned *et al.* (2016). Predictions were made for every segment of the digital river network using random forest models fitted to between 354 sites (TN) and 586 sites (NO3N). Assessment of model performance indicated these predictions were accurate and with low bias (Larned *et al.*, 2016).

Target NO3N and DRP concentrations for periphyton

Nutrient concentration criteria for TN and DRP that were defined to achieve objectives corresponding to the NOF periphyton A, B and C bands were obtained from Snelder (2018). These concentration criteria were derived using models fitted to periphyton and nutrient observations made at NRWQN sites over the 22-year period from 1989 to 2010. The derived concentration criteria were tested using independent data obtained from 173 regional council monitoring sites that were well distributed nationally. The tests indicated the criteria were too restrictive for TN at the test sites. The criteria were therefore recalibrated to better agree with the test sites and those recalibrated criteria are used in this study. The criteria are designed to achieve a specified NOF periphyton objective at a proportion of locations (referred to as the spatial exceedance criteria by Snelder (2018). In this study, the TN and DRP criteria are associated with the periphyton C/D band and the 10% spatial exceedance criteria was used. TN was converted to equivalent NO3N based on the observed ratio of NO3N to TN at the 450 sites with MCI and nutrient data.

NOF NO3N and DRP criteria

The NOF NO3N and DRP criteria for every segment of the digital river network was derived by first obtaining NO3N and DRP concentrations to achieve the periphyton bottom of the NOF C band (i.e., the national bottom line) based on the criteria of Snelder (2018). The criteria were compared with estimated reference condition NO3N and DRP concentrations estimated by (McDowell *et al.*, 2013). Where the reference condition estimate was less than the criteria, it was assumed that a realistic C band estimate could not be made.

An index representing the size of bed substrate material was obtained from FENZ. The assumption of Snelder *et al.* (2013) was followed that network segments with index values of three or less have fine substrates ('soft bottoms') and do not support conspicuous periphyton. In these cases, it was assumed that the NOF nitrate-nitrogen toxicity attribute would apply. The NOF C band (i.e., the national bottom line) for nitrate-nitrogen toxicity of 6.9 mg/l was used to set the NOF NO3N criterion for these segments. This may over-estimate the number of network segments where the NOF nitrate-nitrogen toxicity attribute would apply. Where soft-bottom rivers discharge into sensitive lakes or estuaries, the NOF requires that nitrogen and phosphorus loads and/or concentrations are set to achieve trophic objectives in the downstream receiving environment. The extent to which nitrogen and phosphorus loads and/or concentrations in rivers will need to be managed to achieve trophic state objectives in sensitive downstream receiving environments has not been evaluated. The spatial assessment presented here therefore represents the maximum levels of NO3N and DRP that are permissible under the NOF.

Methods

Assessment of relationship of NO3N and DRP with MCI within river classes

Linear regression was used to quantify the relationship of NO3N and DRP with MCI for the whole dataset of 450 sites with MCI and nutrient data (national scale) and within REC Source-of-flow and FWENZ classes.

All monitoring sites were allocated to REC Source-of-flow classes based on their location on the digital river network. Some Source-of-flow classes had poor representation within the monitoring network (i.e., < 10). REC classes with poor representation were aggregated into the class that was closest in environmental terms so that all analysed classes had at least 10 representative sites (Table 3).

Table 3. REC Source-of-flow classes used in the analysis

Source-of-flow class	Comprises	Description – catchment dominated by:
WW/L	WW/L, WW/H, WW/Lk, WX/L	Warm wet lowlands, including extremely wet lowland and lake-fed
WD/L	WD/L, WD/Lk	Warm dry lowlands, including lake-fed
CW/H	CW/H	Cool wet hills
CX/H	CX/H, CX/M	Cool extremely wet hills and mountains
CW/L	CW/L, CW/Lk	Cool wet lowland and lake fed
CD/H	CD/H, CD/M	Cool dry hills and mountains
CW/M	CW/M	Cool wet mountains
CD/L	CD/L, CD/Lk	Cool dry lowlands and lakes
CX/L	CX/L, CW/Lk	Cool extremely wet lowlands and lake fed

All monitoring sites were allocated to FWENZ level-one classes based on their location on the digital river network. Only seven classes were represented within the monitoring network (Table 4**Error! Reference source not found.**). Two classes (B and I) had only one site and these sites were removed from the analysis.

Table 4. FWENZ level-one classes used in the analysis.

FWENZ level one class	Class description (from Leathwick *et al.*, (2011)
A	Low-elevation rivers and streams
C	Lowland hill country gravel-bed streams
D	South Island low-elevation streams and rivers in dry inland areas
E	South Island low-elevation large rivers in dry, inland areas
G	Mid-elevation streams and rivers in dry inland areas

The significance of the regressions was tested using the non-parametric Kendall's Tau correlation statistic. The linear regressions were used to estimate the concentrations of NO3N and DRP that were associated with MCI scores of 90, 100 and 120. These scores provided benchmarks of ecosystem health that are approximately comparable with NOF A, B and C bands and enabled the analysis to indicate whether the NO3N and DRP associated with ecosystem health differed between classes.

Variance partitioning analysis

An assessment of the 'uniqueness' of the of the ecosystem health NO3N/DRP relationship was conducted based on variance partitioning. The analysis acknowledges that environmental variables other than NO3N/DRP may influence biological communities, including the catchment and river network characteristics shown in Table 2. Many of these variables are correlated because they share strong hierarchical relationships and because they tend to vary monotonically as a function of position in the river network (e.g., Montgomery, 1999; Poff, 1997; Vannote *et al.*, 1980). Correlation between these environmental variables may lead to overestimating the strength of relationships between NO3N/DRP and biological characteristics if covariance is not taken into account (Borcard *et al.*, 1992; Fortin and Dale, 2005; Legendre and Legendre, 1998).

Variance partitioning analysis was used to quantify the strength of relationships between NO3N/DRP and MCI while considering the extent to which these relations may be overestimated if other environmental variables are not considered. The analysis used a procedure that is based on multiple linear regression to partition the total explained variation in MCI scores into 15 components that included the individual, shared and unique contributions of the three sets of variables representing the factors: nutrient concentrations, catchment variables and segment variables (Borcard *et al.*, 1992). The significance of all components was tested using permutation tests. The significance of the unique fractions was tested by running the other set of variables as co-variables (i.e., their effect was removed; Legendre and Legendre, 1998).

The variance partitioning analysis was performed using the national dataset (i.e., regional council and NRWQN sites) and just the NRWQN dataset. General agreement between these two sets of results is interpreted as evidence that the analyses are not compromised by use of data from multiple collection agencies.

Estimates of explained variation derived from samples are generally biased (Zar, 1999). This bias is influenced by the number of independent variables in the model and sample size. The method of Peres-Neto *et al.* (2006) was used to adjust the estimate of variation explained by each set of variables to make valid comparisons between sets of variables of differing size. All analyses and variance partitioning were performed in R using the 'vegan' package (R Development Core Team, 2004).

Spatial assessment of impact of the DIN and DRP bottom lines

A spatial assessment of the impact of the DIN and DRP bottom lines was performed for both the estimated NOF DIN and DRP criteria and the thresholds proposed by Death *et al.* (2018). The spatial assessment evaluated the following:

1. The locations where current concentrations exceed the NOF DIN and DRP criteria.
2. The locations where current concentrations exceed Death *et al.'s (*2018) proposed thresholds.
3. The locations where current concentrations exceed Death *et al.'s (*2018) proposed thresholds but not the NOF DIN and DRP criteria.
4. The amount by which concentrations that currently exceed Death *et al.'s (*2018) proposed thresholds, but comply with the NOF criteria, would need to decrease to comply with Death *et al.'s (*2018) proposed thresholds.
5. The locations where the NOF DIN and DRP criteria are less than Death *et al.'s (*2018) proposed thresholds.

Results

Relationships of NO3N and DRP with MCI at the national scale

The relationships of $\log_{10}$ transformed NO3N and DRP with MCI at the national scale are shown on Figure 3. NO3N and DRP explained 14% and 8% of the variation in site median MCI scores, respectively (Table 5). The concentrations of NO3N and DRP associated with the nominated MCI scores of 90, 100 and 120 were substantially higher than Death *et al.'s* (2018) proposed thresholds except for NO3N for MCI of 100 and 120 (Table 5).

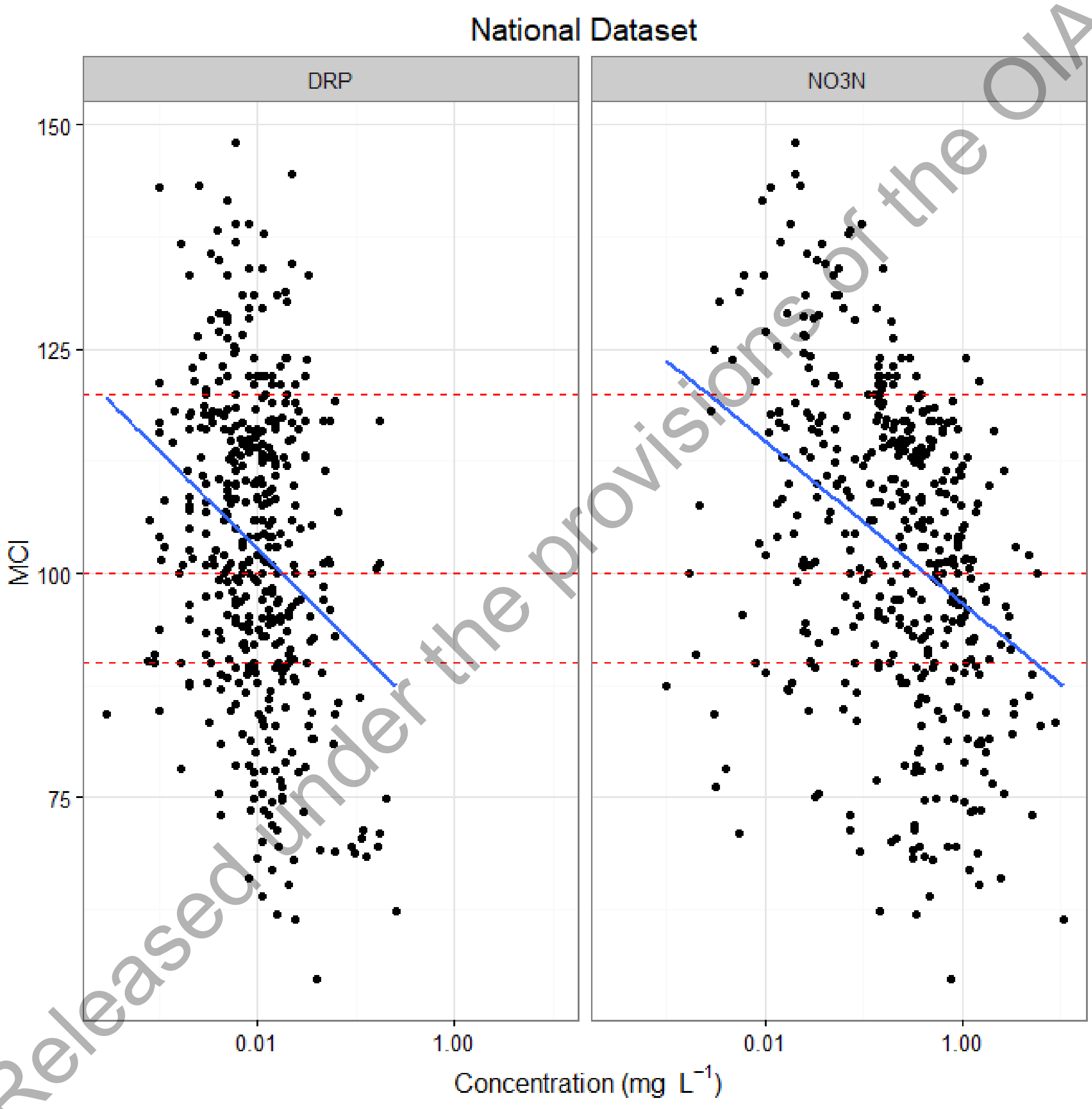


Figure 3. Relationships of site median values of NO3N and DRP with site median MCI scores at national scale. The blue lines represent regressions of the MCI against the nutrient concentration. The red lines indicate nominated MCI scores of 90, 100 and 120.

Table 5. Results of linear regression of MCI against the nutrient concentration at the national scale. Significant p-values are in bold. The concentrations of DIN and DRP that were associated with MCI scores of 90, 100 and 120 are shown and the corresponding Death et al. (2018) thresholds are shown in parentheses.

Nutrient	N	Coefficient	R^2	p-value	Kendall Tau	Kendall p-value	MCI 90	MCI 100	MCI 120
DRP	445	-10.98	8	**<0.001**	-0.17	**<0.001**	0.14 (0.038)	0.02 (0.019)	0 (0.006)
NO3N	445	-9.02	14	**<0.001**	-0.29	**<0.001**	5.5 (0.89)	0.4 (0.46)	0.003 (0.1)

Relationship of NO3N and DRP with MCI in REC classes

The relationships of NO3N and DRP with MCI within the REC classes are shown in Figure 4 and Figure 5. NO3N and DRP did not have significant relationships with MCI scores in many classes (Table 6 and Table 7). The concentrations of NO3N and DRP associated with the nominated MCI scores of 90, 100 and 120 differed substantially between classes (Table 6 and Table 7Table 7), but could not be estimated for some classes because most of the MCI scores in those classes were either lower (e.g., class CD/L; Figure 4 and Figure 5) or higher (e.g., class CX/H; Figure 5) than the nominated values within the class.

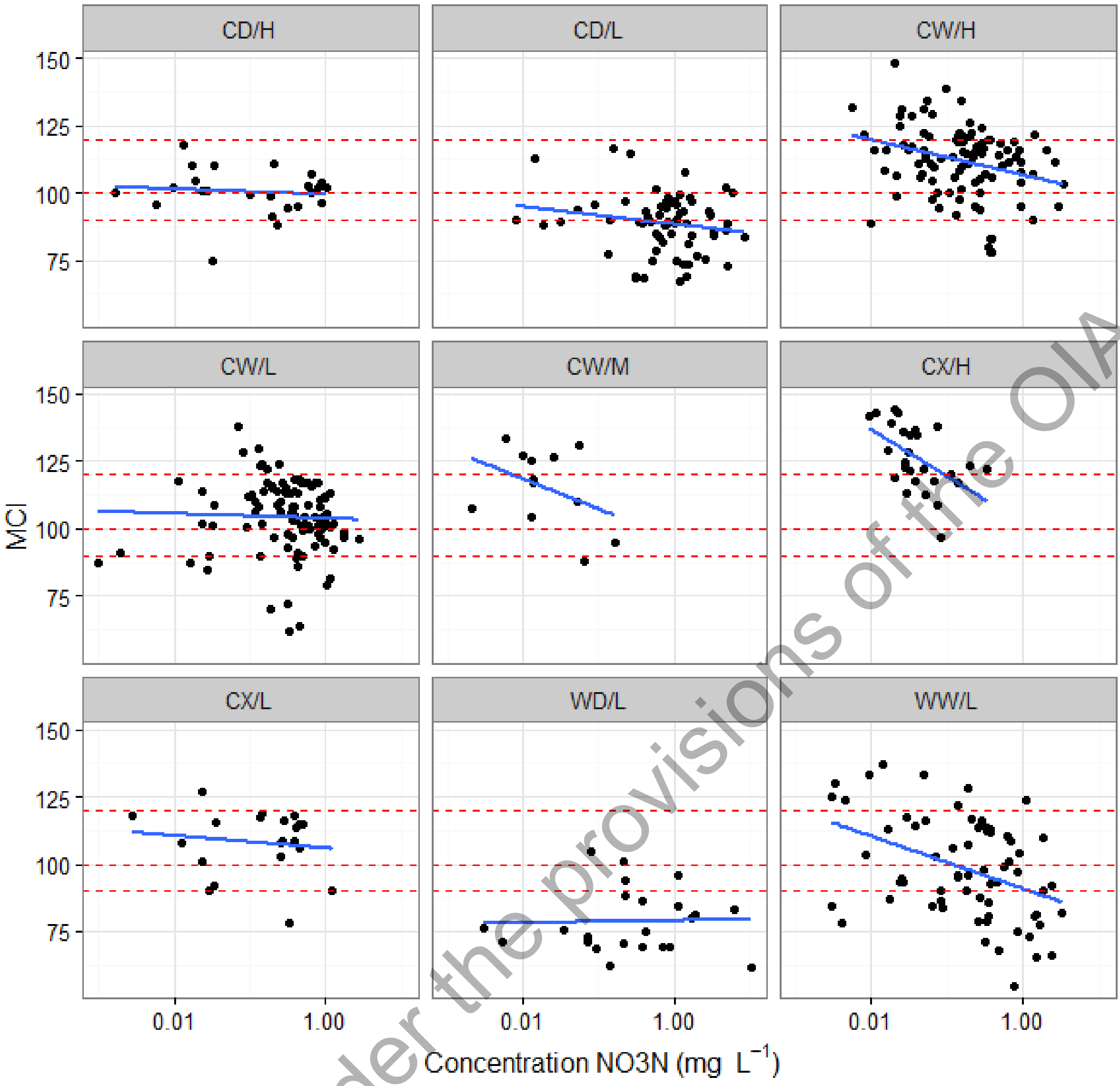


Figure 4. Relationships of site median values of NO3N with site median MCI scores in REC Source-of-flow classes. The blue lines represent a regression of the MCI against the nutrient concentration. The red lines indicate nominated MCI scores of 90, 100 and 120.

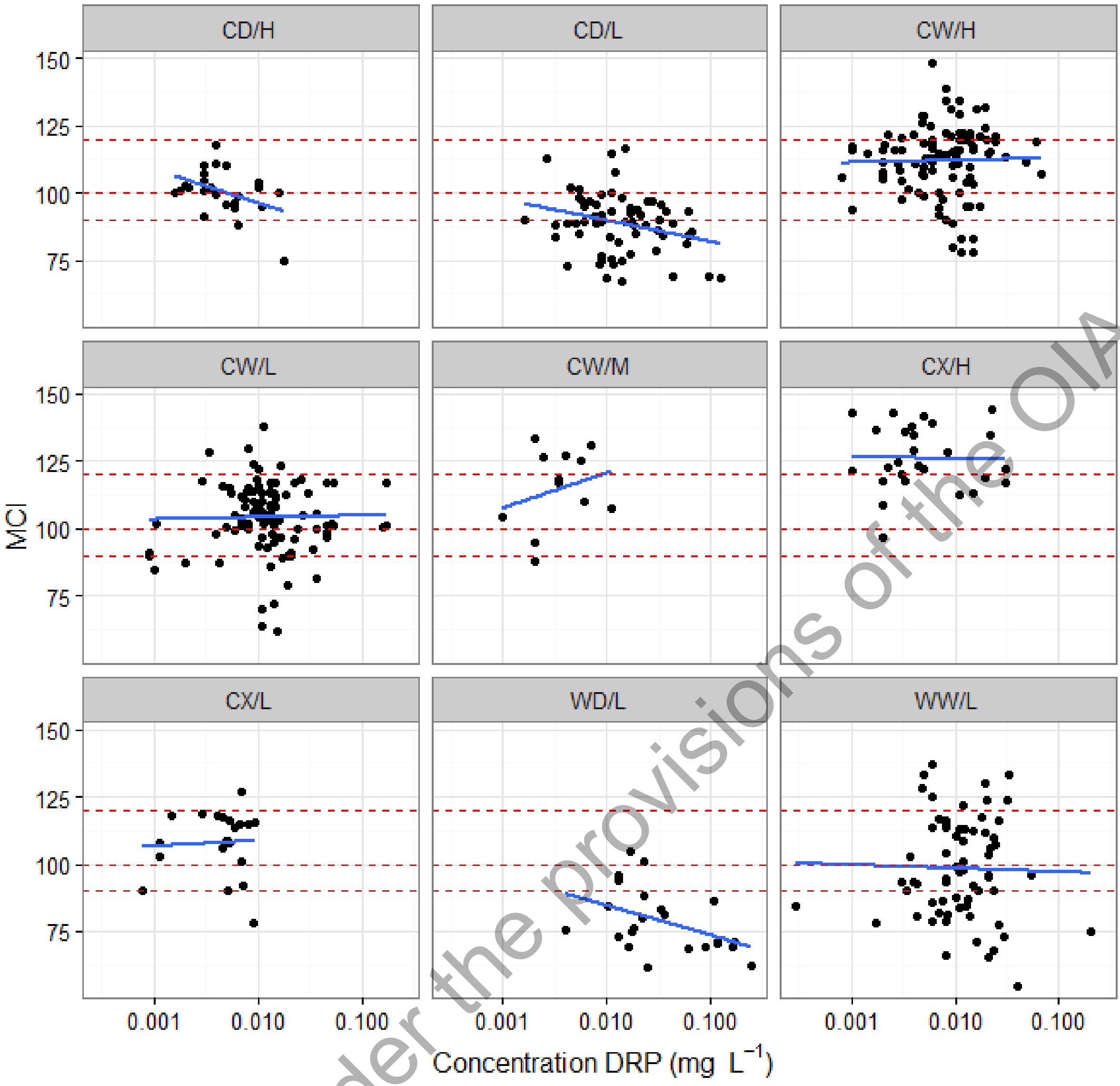


Figure 5. Relationships of site median values of NO3N with site median MCI scores in REC Source-of-flow classes. The blue lines represent a regression of the MCI against the nutrient concentration. The red lines indicate nominated MCI scores of 90, 100 and 120.

Table 6. Results of linear regression of MCI against the NO3N concentration within REC Source-of-flow classes. NA values indicate the concentration associated with the nominated MCI value could not be estimated. Significant p-values are in bold.

Nutrient	N	Coefficient	R^2	p-value	Kendall Tau	Kendall p-value	MCI 90	MCI 100	MCI 120
CD/H	24	-0.98	-4	0.66	-0.06	0.69	NA	0.204	NA
CD/L	68	-3.27	2	0.14	-0.07	0.41	0.29	0	NA
CW/H	101	-6.54	8	**0**	-0.2	**0.003**	NA	9.951	0.009
CW/L	102	-0.88	-1	0.7	-0.15	**0.02**	NA	NA	NA
CW/M	13	-11.2	9	0.17	-0.3	0.15	3.4	0.439	0.007
CX/H	27	-17.31	24	**0.01**	-0.41	**0.003**	4.9	1.306	0.091
CX/L	21	-2.32	-3	0.57	-0.06	0.69	NA	NA	NA
WD/L	23	0.45	-5	0.89	-0.03	0.85	1.2	NA	NA
WW/L	66	-9.72	16	0	-0.27	0.001	13.215	0.116	0.001

Table 7. Results of linear regression of MCI against the DRP concentration within REC Source-of-flow classes. NA values indicate the concentration associated with the nominated MCI value could not be estimated. Significant p-values are in bold.

Nutrient	N	Coefficient	R^2	p-value	Kendall Tau	Kendall p-value	MCI 90	MCI 100	MCI 120
CD/H	24	-12.41	14	**0.04**	-0.23	0.13	0.03	0.005	0
CD/L	68	-7.81	6	**0.02**	-0.16	**0.05**	0.009	0	NA
CW/H	101	0.86	-1	0.81	0.04	0.54	NA	NA	NA
CW/L	102	0.71	-1	0.83	-0.06	0.40	NA	NA	NA
CW/M	13	13.04	-1	0.38	0.09	0.66	NA	0	0.009
CX/H	27	-0.47	-4	0.93	0	0.98	NA	NA	NA
CX/L	21	2.06	-5	0.81	-0.05	0.73	NA	NA	NA
WD/L	23	-11.03	16	**0.04**	-0.31	**0.04**	0.003	0	NA
WW/L	66	-1.27	-1	0.83	-0.01	0.91	NA	0.001	NA

Relationship of NO3N and DRP with MCI in FWENZ classes

The relationships of NO3N and DRP with MCI within the FWENZ classes are shown on Figure 6 and Figure 7. NO3N had significant relationships with MCI scores in classes C and G but did not have a significant relationship in class A despite being represented by 70 sites (Table 8). DRP had significant relationships with MCI scores in classes A, C and G (Table 9). There were few representative sites in classes D and E (Table 8 and Table 9). The concentrations of NO3N and DRP associated with the nominated MCI scores of 90, 100 and 120 differed substantially between FWENZ classes (Table 8 and Table 9) but could not be estimated for some classes because MCI scores were substantially lower than the nominated values within the class (e.g., class A; Figure 6 and Figure 7).

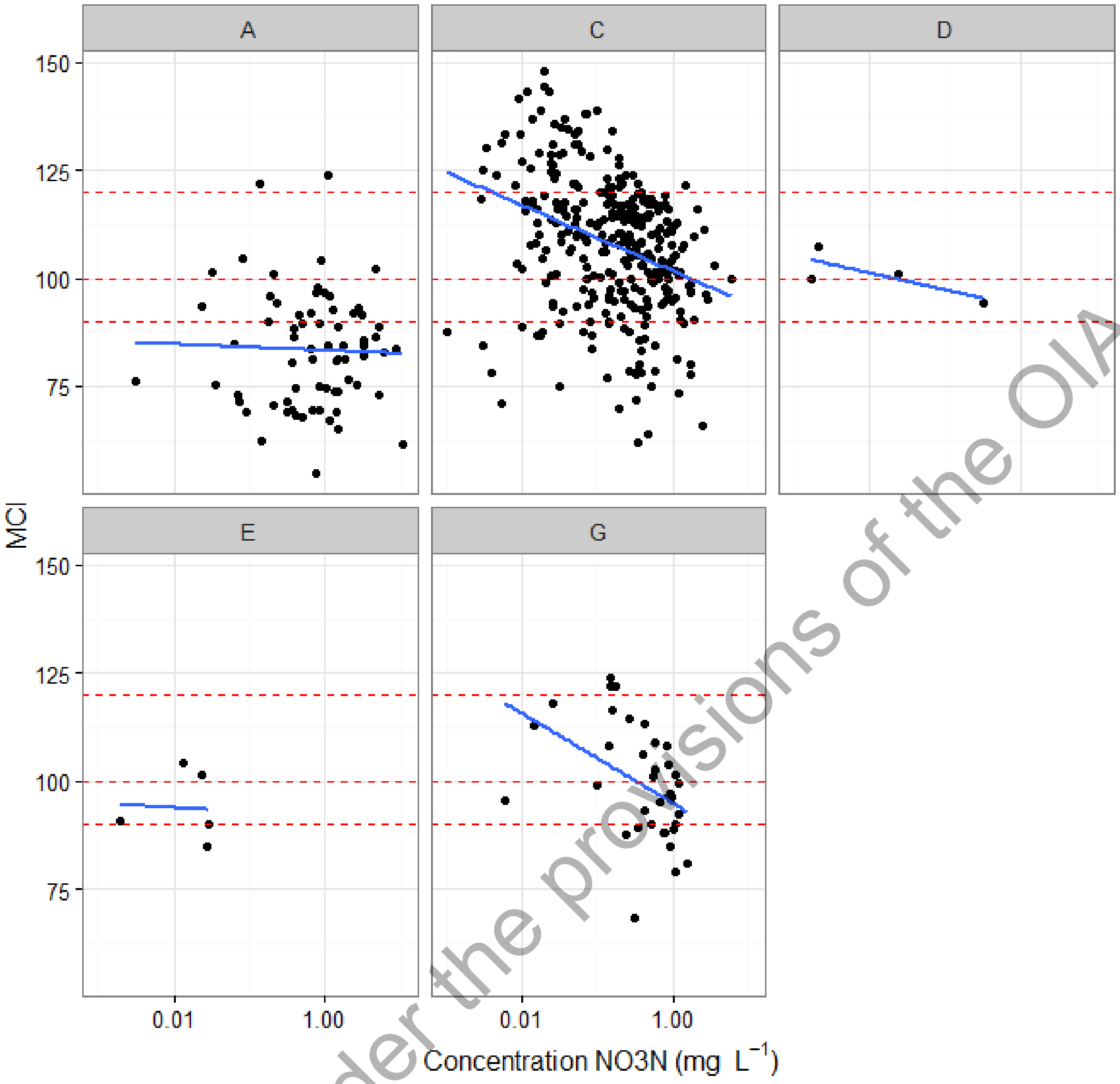


Figure 6. Plot showing relationships of site median values of NO3N with site median MCI scores in FWENZ classes. The blue lines represent a regression of the MCI against the nutrient concentration. The red lines indicate nominated MCI scores of 90, 100 and 120.

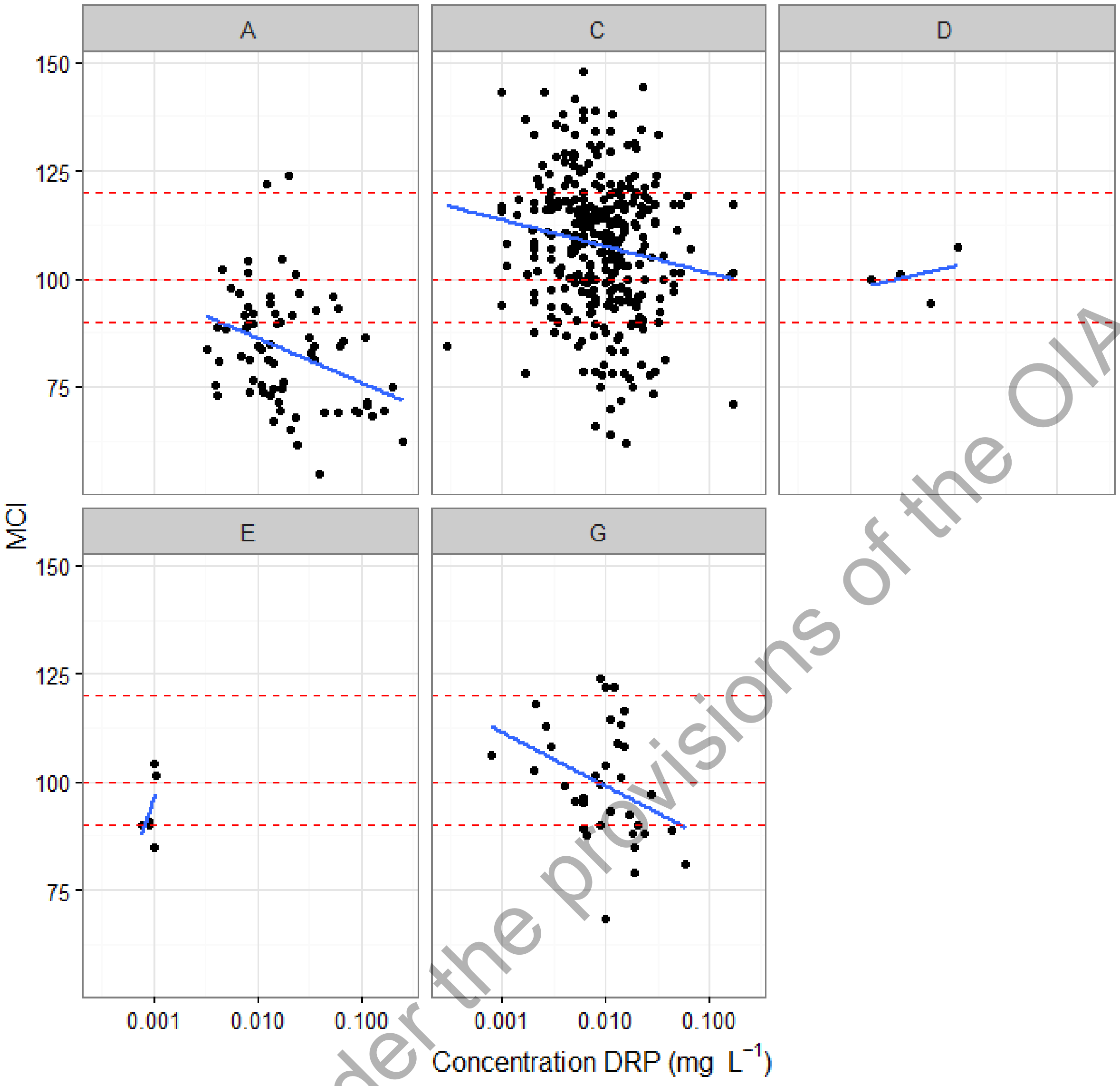


Figure 7. Plot showing relationships of site median values of DRP with site median MCI scores in FWENZ classes. The blue lines represent a regression of the MCI against the nutrient concentration. The red lines indicate nominated MCI scores of 90, 100 and 120.

Table 8. Results of linear regression of MCI against the NO3N concentration within FWENZ classes. NA values indicate the concentration associated with the nominated MCI value could not be estimated. Significant p-values are in bold.

Nutrient	**N**	**Coefficient**	**R^2**	**p-value**	**Kendall Tau**	**Kendall p-value**	**MCI 90**	**MCI 100**	**MCI 120**
A	70	-0.7	-1	0.8	-0.01	0.883	NA	NA	NA
C	326	-7.6	11	**0**	-0.25	**0**	34.2	1.67	0.004
D	4	-3.9	42	0.2	-0.33	0.7	7.2	0.02	NA
E	6	-0.96	-25	0.9	-0.47	0.3	85.3	NA	NA
G	37	-10.4	16	**0.01**	-0.34	**0.003**	2.9	0.32	0.004

Table 9. Results of linear regression of MCI against the DRP concentration within FWENZ classes. NA values indicate the concentration associated with the nominated MCI value could not be estimated. Significant p-values are in bold.

Nutrient	N	Coefficient	R^2	p-value	Kendall Tau	Kendall p-value	MCI 90	MCI 100	MCI 120
A	70	-10.3	11	**0**	-0.24	**0.004**	0.004	0	NA
C	326	-6.1	2	**0**	-0.11	**0.003**	6.9	0.16	NA
D	4	5.4	-30	0.63	0.33	0.75	NA	0.003	15.5
E	6	70.7	1	0.36	0.21	0.56	0.001	0.001	0.002
G	37	-12.5	11	**0.03**	-0.28	**0.016**	0.05	0.008	0

Variance partitioning

National dataset

The results of the variance partitioning analysis performed using the national dataset (i.e., regional council and NRWQN sites) are shown graphically on Figure 8. The outer box represents the total variation in the site MCI scores. The Venn-diagram within the box represents the total explained variation (r^2 = 0.47) in MCI scores (residual unexplained variation = 0.53). The variation explained by each group of variables is represented by the sum of the values lying within each of the three circles that represent the catchment, segment and nutrient variable groups. For catchment and segment variable groups, see Table 2. The unique contribution of each of the groups of variables is shown by the parts of the circles that do not overlap with the other circles (labelled [a], [b] and [c]). The explained variation that is shared is shown by the intersection areas of the three circles (labelled [d], [e] [f] and [g]).

Nutrients (i.e., the combination of NO3N and DRP) individually explained (i.e., A+D +F+G in Figure 8) 18% of the variation. The segment and catchment variables individually explained 15% and 40% of the variation respectively. Nutrients uniquely explained (i.e., A in Figure 8) 1% of the variation in MCI scores, that is, after accounting for variation explained by the segment and catchment variables, nutrients explained a further 1% of the variation. Segment and catchment uniquely explained 6% and 19% of the variation respectively.

Two further analyses were performed to assess the variation in MCI explained by nutrients after controlling for segment and for catchment. Nutrients explained 13% of the variation in MCI after controlling for the segment variables but only explained 1% of the variation after controlling for the catchment variables.

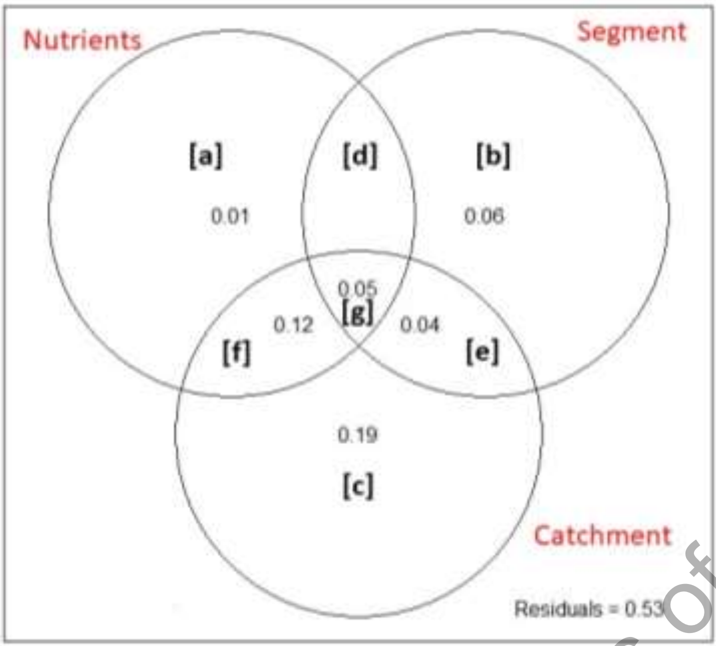


Figure 8. Schematic diagram of all components of variation provided by the variance partitioning.

Table 10. Components of explained variation provided by the variance partitioning. Refer to Figure 8 for the component labels a, b, c etc.

Component	Type of contribution	Adjusted variation explained (r^2)	P-value
[a+d+f+g] = Nutrients	Nutrients individual	0.18	<0.001
[b+d+e+g] = Segment	Segment individual	0.15	<0.001
[c+e+f+g] =Catchment	Catchment individual	0.40	<0.001
[a+b+d+e+f+g] = Nutrients + Segment	Joint	0.28	<0.001
[a+c+d+e+f+g] = Nutrients + Catchment	Joint	0.41	<0.001
[b+c+d+e+f+g] = Segment + Catchment	Joint	0.46	<0.001
[a+b+c+d+e+f+g] = All	Joint	0.47	<0.001
[a] = Nutrients | Segment + Catchment	Nutrients unique	0.01	<0.001
[b] = Segment | Nutrients + Catchment	Segment unique	0.06	<0.001
[c] = Catchment | Nutrients + Segment	Catchment unique	0.19	<0.001
[d]	Shared	0.00	Not testable
[e]	Shared	0.04	Not testable
[f]	Shared	0.12	Not testable
[g]	Shared	0.05	Not testable
[h]	Shared	0.53	Not testable
[a]+[f] = Nutrients | Segment	Nutrients unique controlling for segment	0.13	<0.001
[a]+[d] = Nutrients | Catchment	Nutrients unique controlling for catchment	0.01	0.005

The results of the variance partitioning analysis performed using the NRWQN dataset are shown graphically on Figure 9. The model explained a total variation in MCI scores (r^2) of 0.68 (residual unexplained variation = 0.32). The analysis indicated that nutrients (i.e., the combination of NO3N and DRP) individually explained (i.e., A+D+F+G in Figure 9) 2% of the variation. The segment and catchment variables individually explained 30% and 44% of the variation respectively. Nutrients uniquely explained (i.e., A in Figure 7) 1.5% of the variation in MCI scores, that is, after accounting for variation explained by the segment and catchment variables, nutrients explained a further 1.5% of the variation. Segment and catchment uniquely explained 24% and 39% of the variation respectively.

Permutation tests indicated that the component of variation that was uniquely explained by nutrients was not statistically significant (p=0.19). The other components of variation that were testable were all significant (p<0.001).

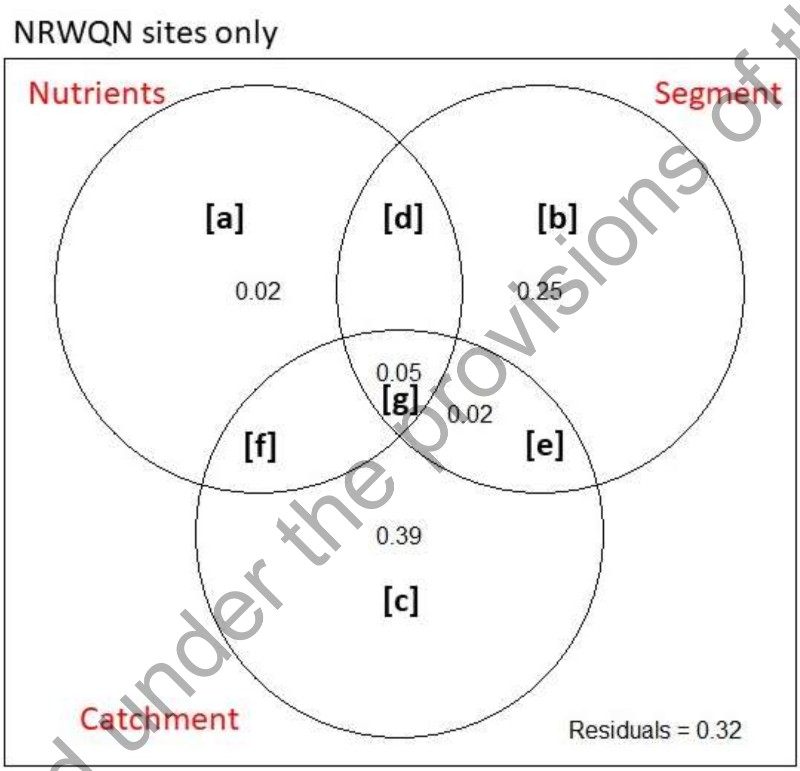


Figure 9. Schematic diagram of all components of variation provided by the variance partitioning performed using only the NRWQN sites.

Spatial assessment of impact of proposed bottom lines for DIN and DRP

The ratio of median NO3N to median TN was variable over the monitoring sites and ranged from 0.3 to >1 (Figure 10). The median of 0.6 was adopted to adjust the periphyton TN criterion, to an equivalent DIN.

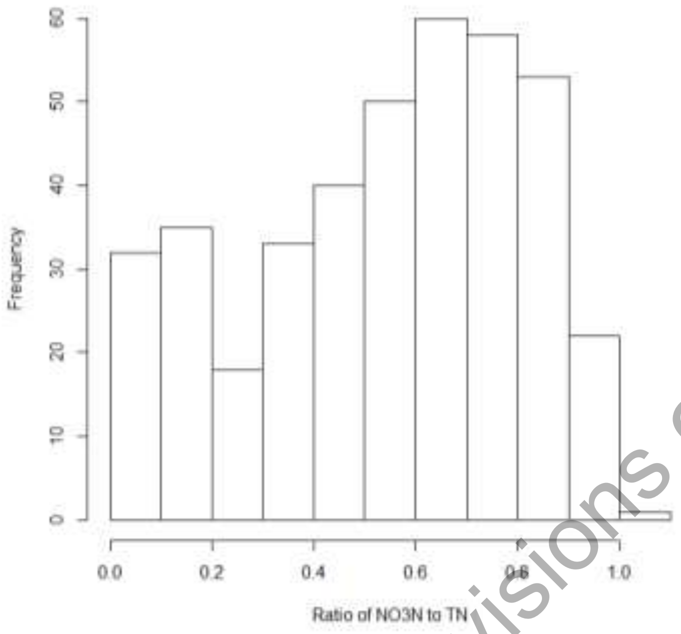


Figure 10. Histogram of the ratio of NO3N to TN at the monitoring sites.

The spatial distribution of the NOF DIN and DRP criteria are shown on Figure 11. The criteria were generally most lenient on the West Coast of the South Island and increased in stringency to the east and north. This reflects the effect on periphyton biomass of reduced high-flow frequency on the east coasts of both islands, and increasing temperature and solar radiation moving north, both of which require reduced DIN and DRP concentrations to achieve a given biomass.

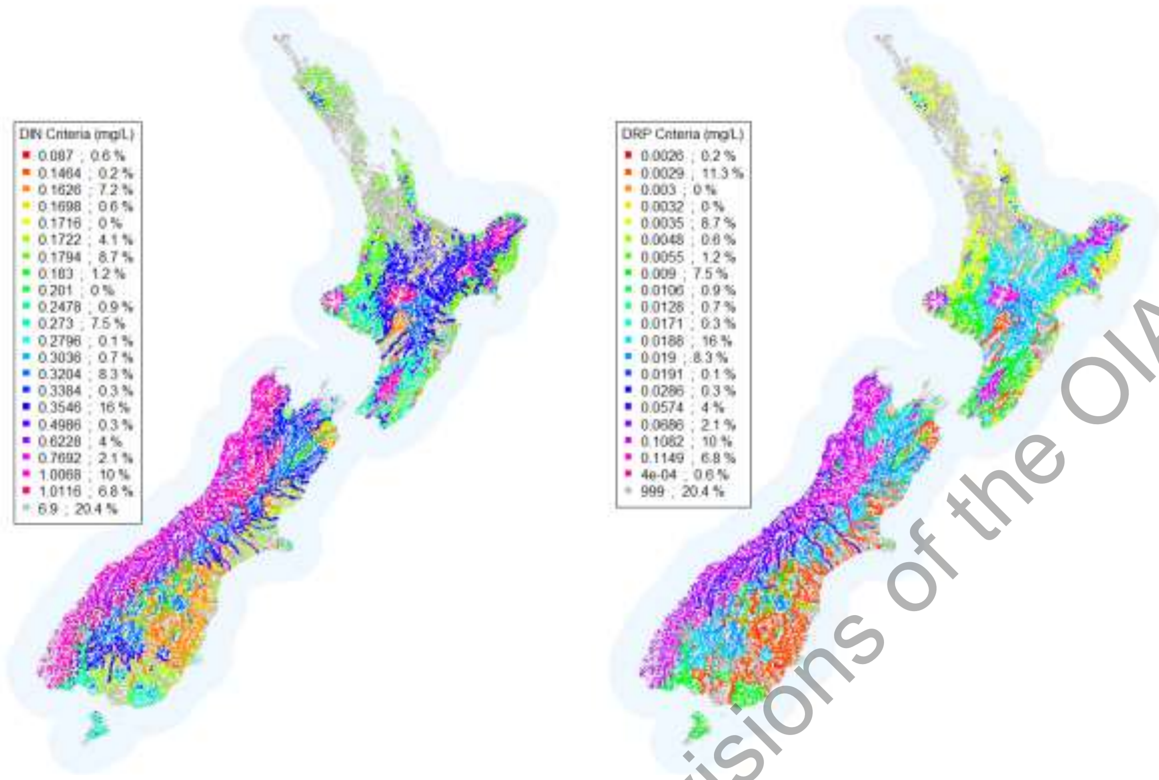


Figure 11. Spatial distribution of NOF DIN and DRP criteria for rivers of order three or greater. Note that rivers with DIN criterion of 6.9mg/l and DRP criterion of 999 mg/l are locations that are estimated to have fine substrates that do not support conspicuous periphyton. The number of significant figures is due to converting the original values that had units of mg/m³.

For rivers of order three or greater, 12% and 15% of segments exceeded the NOF DIN and DRP criteria respectively (Figure 12). When segments of all stream orders were considered, 11% and 16% % of segments exceeded the NOF criteria for DIN and DRP respectively. These rivers represent locations at which the NOF requires that current DIN concentrations are decreased and/or alternative methods for controlling periphyton biomass are implemented. A high proportion of rivers had estimated reference state for DRP that was greater than the estimated NOF criterion and these rivers are not shown on the map in Figure 12. This probably reflects uncertainties associated with estimating both the reference condition and the DRP criterion. The analysis was unable to determine whether current DRP concentrations need to decrease at these locations to comply with the NOF.

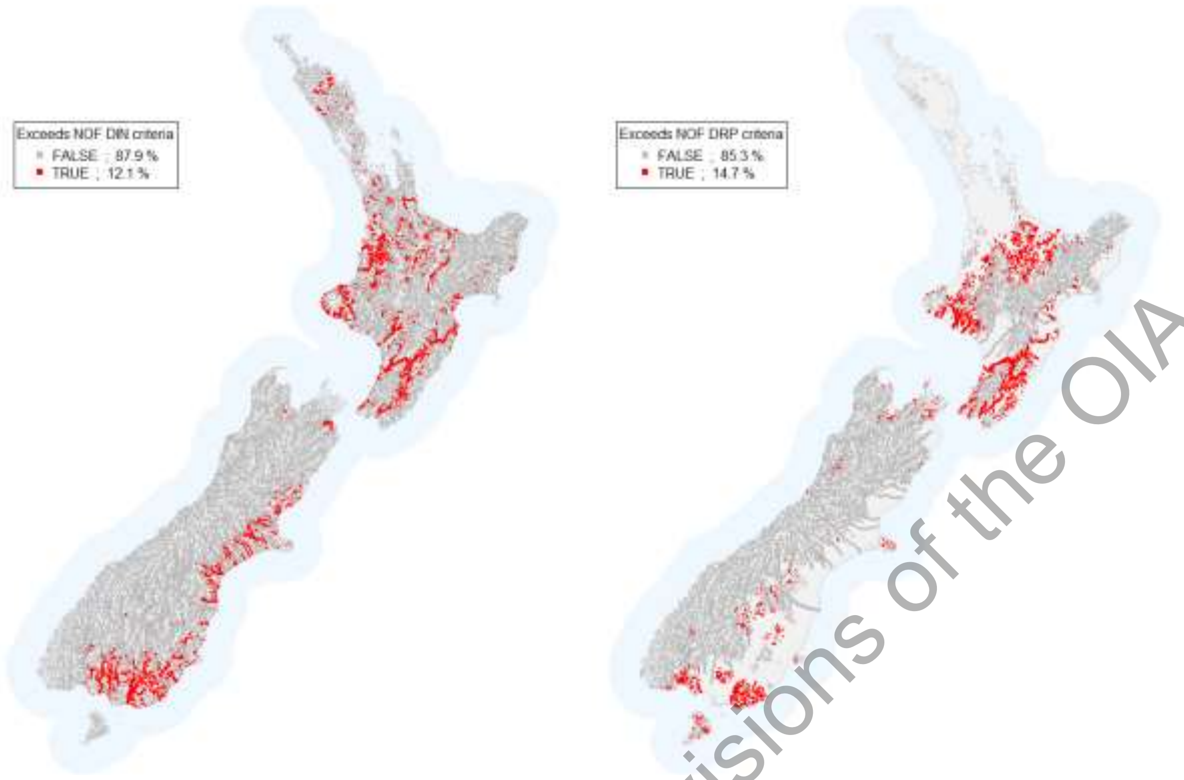


Figure 12. Spatial distribution of rivers of order three or greater for which DIN and DRP are assessed to exceed NOF criteria. Rivers not shown on the DRP map had reference state for DRP that was greater than the estimated NOF criterion.

For rivers of order three or greater, 6% and 1% of segments exceeded the Death *et al.'s* (2018) proposed bottom line thresholds for DIN and DRP respectively (Figure 13). When segments of all stream orders were considered, 9% and 0.6% of segments exceeded Death *et al.'s* (2018) proposed bottom line thresholds for DIN and DRP respectively.

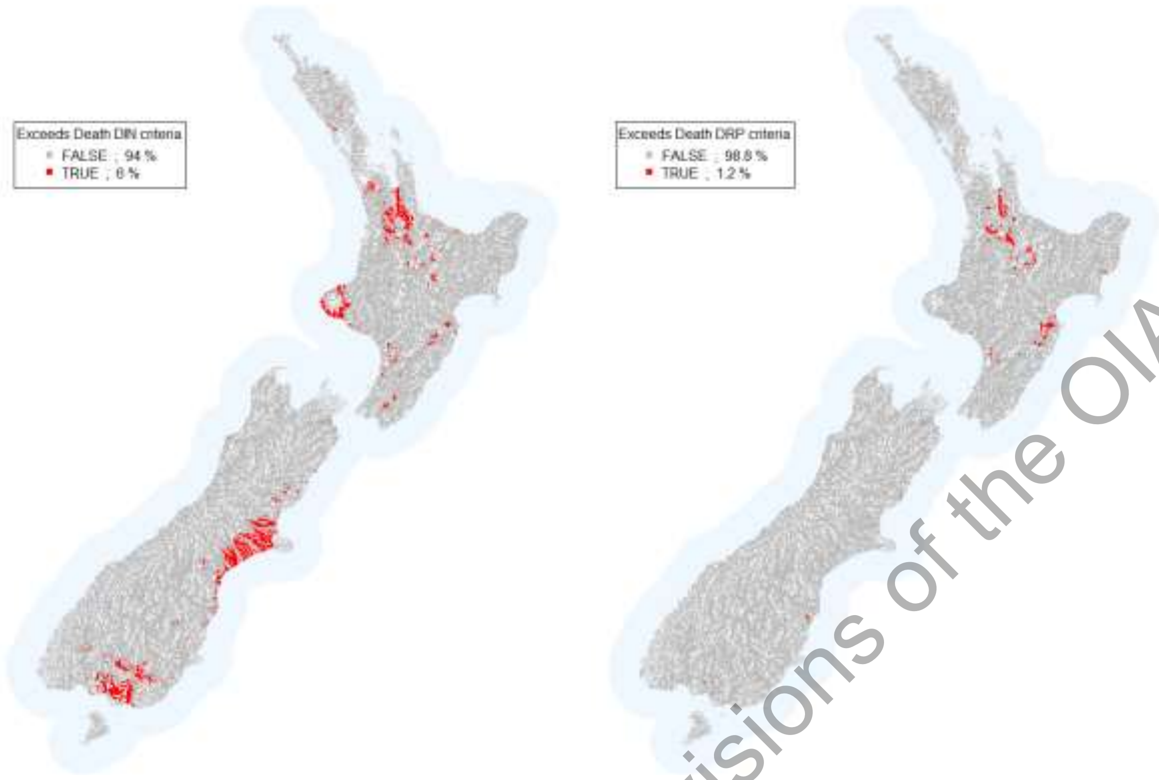


Figure 13. Spatial distribution of rivers of order three or greater for which DIN and DRP are predicted to exceed the Death et al.'s (2018) proposed bottom line thresholds.

For 4.5% of segments of order three or greater, Death *et al.'s* (2018) proposed DIN threshold was exceeded but the NOF DIN criterion was not exceeded (Figure 14). When segments of all stream orders were considered, 7% of segments exceeded Death *et al.'s* (2018) proposed DIN threshold but the NOF DIN criterion was not exceeded. It is noted that where streams discharge into sensitive lakes or estuaries the NOF requires that nitrogen and phosphorus are managed to achieve acceptable trophic states in the downstream receiving environment. This may have the effect of reducing the NOF DIN and DRP criterion in the streams shown in Figure 14 and reducing the extent of segments that exceed Death *et al.'s* (2018) proposed DIN threshold but do not exceed the NOF DIN criterion. There were no segments for which Death *et al.'s* (2018) proposed DRP threshold was exceeded but the NOF DRP criterion was not exceeded (Figure 14). These rivers represent locations at which there is a risk that DIN is not constraining ecological health based on the demonstration above that there is not a unique relationship between nutrients and MCI.

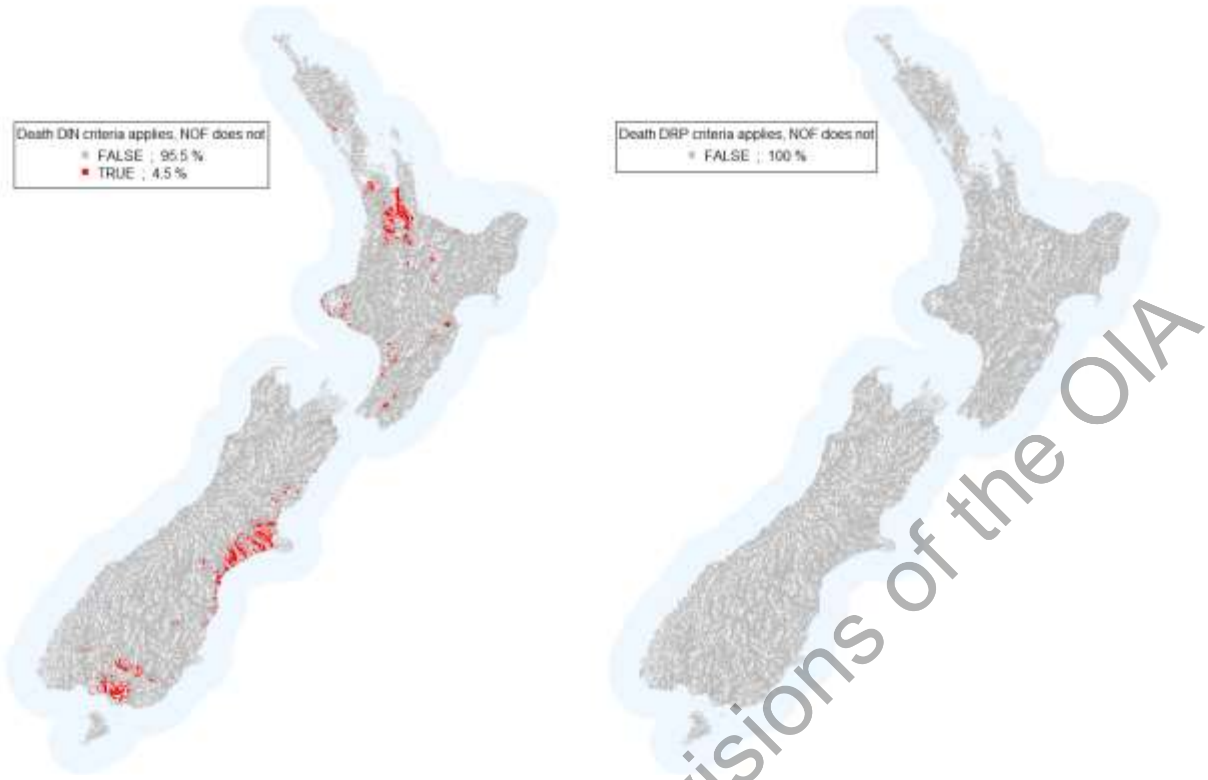


Figure 14. Spatial distribution of rivers of order three or greater for which the Death al.'s (2018) proposed thresholds for DIN and DRP are assessed to be more stringent than the NOF criteria.

The impact of Death *et al.'s* (2018) proposed DIN threshold is demonstrated on Figure 15 as the amount by which concentrations that currently comply with the NOF criterion would need to decrease to comply with Death *et al.'s* (2018) proposed DIN threshold. Note the streams shown in Figure 15 are those shown in Figure 14. Almost all locations (99.9%) shown in both Figure 14 and Figure 15 are streams that are assessed as soft bottomed and for which the worst case scenario would be that the NOF criterion is defined by the nitrate-nitrogen toxicity attribute (because it is assumed the streams do not support conspicuous periphyton).

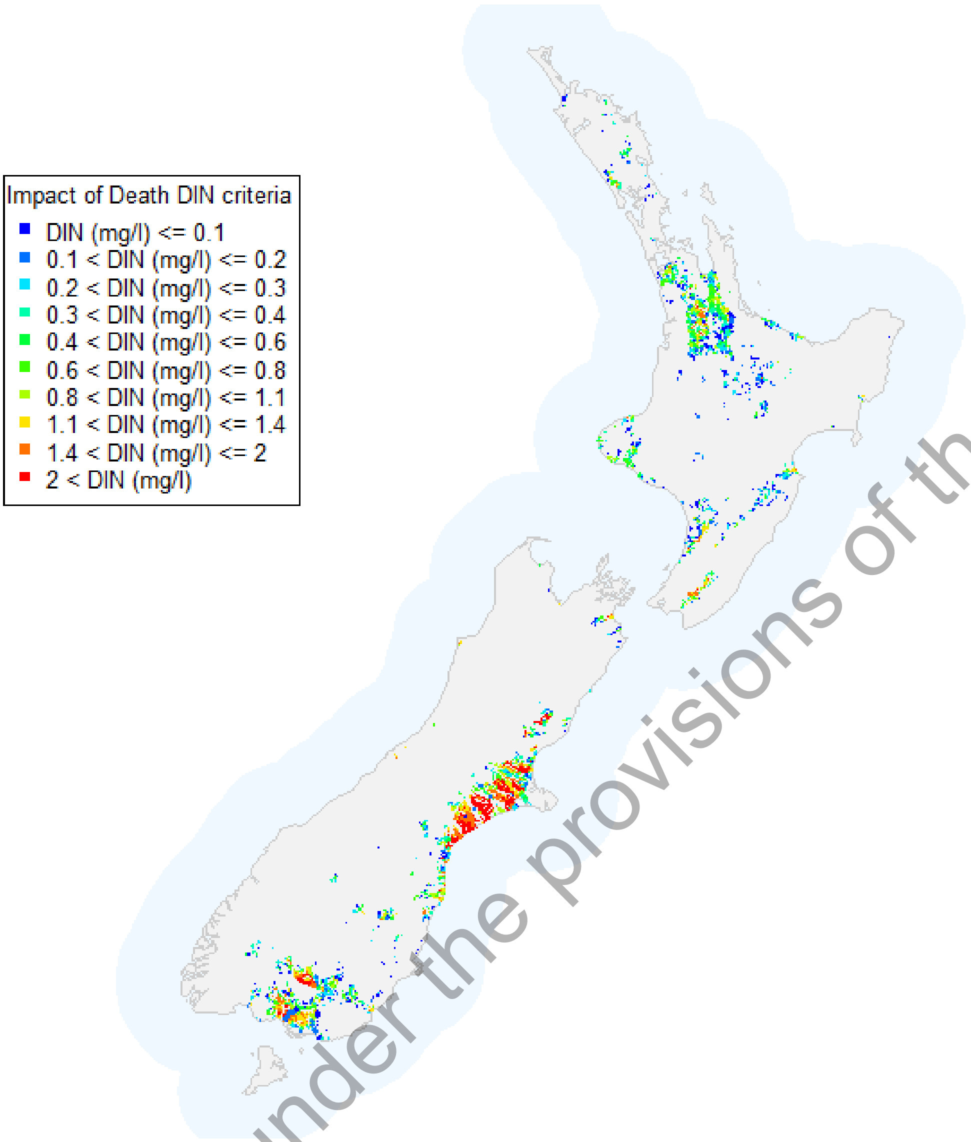


Figure 15. Map showing the amount by which concentrations that currently comply with the NOF criterion would need to decrease to comply with Death et al.'s proposed DIN bottom line threshold. Note the streams shown in the figure are those shown in Figure 14

Locations where the NOF DIN concentration criteria is less (i.e., more stringent) than Death *et al.'s* (2018) proposed DIN threshold are shown on Figure 16. For rivers of order three or greater, the NOF criteria were less than Death *et al.'s* (2018) proposed thresholds for 63% and 57% of segments for DIN and DRP respectively (Figure 16). When segments of all stream orders were considered, the NOF criteria were less than Death *et al.'s* (2018) proposed thresholds for 56% and 52% of segments for DIN and DRP respectively. These rivers represent locations at which there is a risk that regional councils would apply Death *et al.'s* (2018) proposed thresholds without appropriately applying the NOF's DIN and DRP requirements, thereby failing to achieve trophic state objectives (i.e., the more stringent requirement to manage DIN and DRP for trophic state objectives would be ignored). Note that the NOF DIN and DRP requirements include consideration of periphyton and trophic objectives in downstream receiving environments.

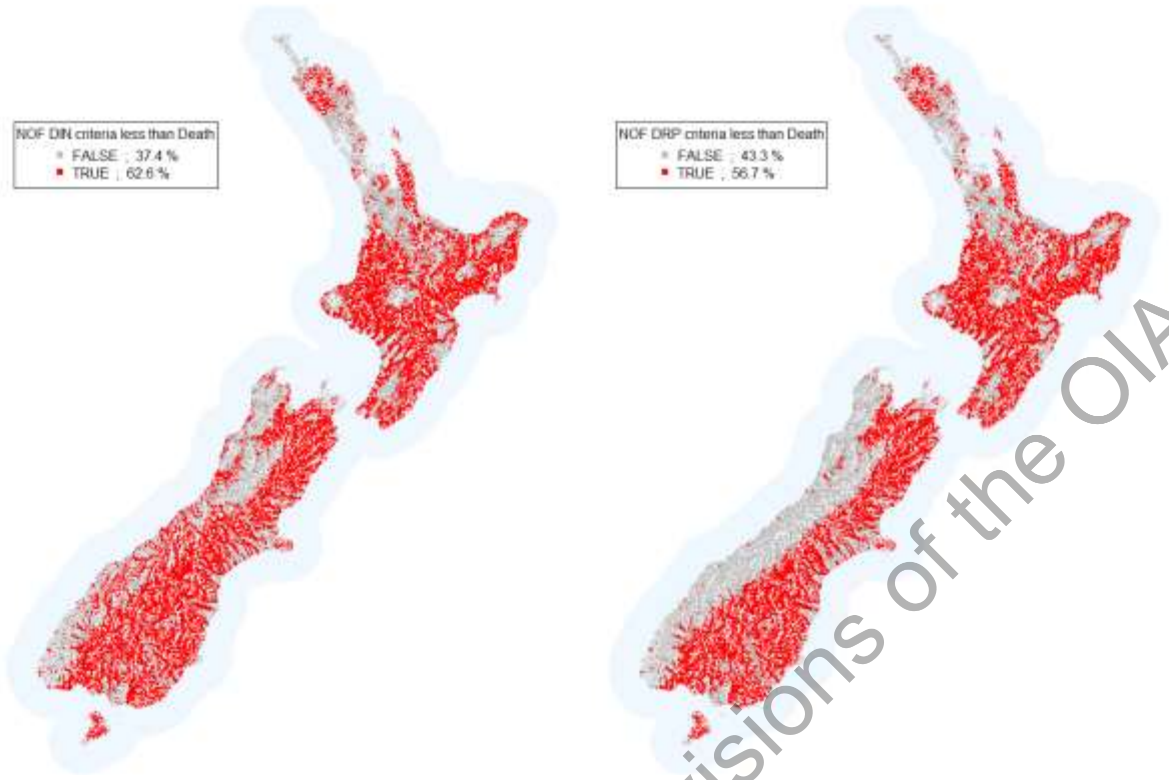


Figure 16. Locations where the NOF criterion is less than the Death et al.'s (2018) proposed thresholds.

Summary

Relationship of NO3N and DRP at scales smaller than national

The analysis of the relationship of NO3N and DRP with MCI within REC and FWENZ classes indicated there was no statistically significant relationship for some classes of both classifications. In addition, when there were relationships, the levels of NO3N and DRP associated with the nominated MCI values of 90, 100 and 120 differed between some classes.

The first result suggests that there is a lack of evidence that variation in MCI in some classes is associated with nutrient concentrations. It is noted that site median values of NO3N and DRP varied widely (generally more than 2 orders of magnitude) in all classes. This suggests that each class was reasonably well represented by a gradient in nutrient concentrations and does not suggest that the lack of a relationship was due the class being dominated by highly degraded conditions.

The second result suggests that there is a risk that NO3N and DRP criteria derived from a national scale analysis will be inappropriate in some systems. The risk extends in two directions, the nationally-derived criteria may be under-protective in some systems and over-protective in others.

Variance partitioning

The variance partitioning showed that associations of DIN and DRP with MCI are very weak when the effect of other variables is controlled for. The analysis indicated that MCI and nutrient concentrations co-vary with catchment factors in particular. This indicates that the relationships from which Death *et al.'s* (2018) proposed DIN and DRP thresholds are derived may be confounded by other factors and that the actual cause the observed variation in MCI may be other variables. This does not prove that nutrient concentrations have no effect on MCI scores, however it reduces confidence that manipulating DIN/DRP will bring about improved ecosystem health relationship.

There was general agreement between the variance partitioning analyses performed using the national dataset (i.e., regional council and NRWQN sites) and just the NRWQN dataset. This result is evidence that the analyses are not compromised by use of data from multiple collection agencies.

Spatial assessment

When segments of all stream orders were considered, 9% and 0.6% of segments exceeded Death *et al.'s* (2018) proposed thresholds for DIN and DRP respectively (Figure 14). This contrasts with the finding that 11% and 16% of segments exceeded the NOF criteria for DIN and DRP respectively (Figure 12).

For 4.5% of segments of order three or greater, and for 7% of all segments, Death *et al.'s* (2018) DIN proposed threshold was exceeded, but the NOF DIN criterion was not exceeded (Figure 14). There were no segments for which Death *et al.'s* (2018) proposed DRP threshold was exceeded, but the NOF DRP criterion was not exceeded (Figure 14). Almost all locations (99.9%) where Death *et al.'s* (2018) proposed DIN threshold was exceeded but the NOF DIN criterion was not exceeded are streams that are predicted to be soft bottomed. The NOF criterion for these locations is defined by the nitrate-nitrogen toxicity attribute. At these locations, there is a risk that DIN is not constraining ecological health based on the demonstration above that there is not a unique relationship between nutrients and MCI.

For rivers of all stream orders, the NOF criteria were more stringent than Death *et al.'s* (2018) proposed thresholds in 56% and 52% of segments for DIN and DRP respectively (Figure 14). These rivers represent locations at which there is a risk that regional councils would apply Death *et al.'s* (2018) proposed thresholds without considering the periphyton attribute or the nutrient constraints associated with downstream lakes and estuaries appropriately, thereby failing to achieve trophic state objectives.

References

Borcard, D., P. Legendre, and P. Drapeau, 1992. Partialling out the Spatial Component of Ecological Variation. Ecology 73:1045–1055.

Death, R.G., A. Canning, R. Magierowski, and J. Tonkin, 2018. Why Aren't We Managing Water Quality to Protect Ecological Health? Occasional Report 31. Palmerston North, NZ., p. .

Fortin, M.J. and M.R.T. Dale, 2005. Spatial Analysis – A Guide for Ecologists. Cambridge University Press, Cambridge.

Larned, S., T. Snelder, and M. Unwin, 2016. Water Quality in New Zealand Rivers; Modelled Water Quality State. NIWA CLIENT REPORT, NIWA, Christchurch, New Zealand.

Leathwick, J., T. Snelder, W. Chadderton, J. Elith, K. Julian, and S. Ferrier, 2011. Use of Generalised Dissimilarity Modelling to Improve the Biological Discrimination of River and Stream Classifications. Freshwater Biology 56:21–38.

Legendre, P. and L. Legendre, 1998. Numerical Ecology. Elsevier, Amsterdam, The Netherlands.

McDowell, R.W., T.H. Snelder, N. Cox, D.J. Booker, and R.J. Wilcock, 2013. Establishment of Reference or Baseline Conditions of Chemical Indicators in New Zealand Streams and Rivers Relative to Present Conditions. Marine and Freshwater Research 64:387–400.

Montgomery, D.R., 1999. Process Domains and the River Continuum. Journal of the American Water Resources Association 35:397–410.

Peres-Neto, P.R., P. Legendre, S. Dray, and D. Borcard, 2006. Variation Partitioning of Species Data Matrices: Estimation and Comparison of Fractions. Ecology 87:2614–2625.

R Development Core Team, 2004. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

Snelder, T., 2018. Nutrient Concentration Targets to Achieve Periphyton Biomass Objectives Incorporating Uncertainties. GNS Science Report, Geological and Nuclear Sciences, Wellington, New Zealand.

Snelder, T.H. and B.J.F. Biggs, 2002. Multi-Scale River Environment Classification for Water Resources Management. Journal of the American Water Resources Association 38:1225–1240.

Snelder, T., B. Biggs, C. Kilroy, and D. Booker, 2013. National Objective Framework for Periphyton. NIWA Client Report, Christchurch, New Zealand.

Vannote, R.L., G.W. Minshall, K.W. Cummins, J.R. Sedell, and C.E. Cushing, 1980. The River Continuum Concept. Canadian Journal of Fisheries and Aquatic Sciences 37:130–137.

Wild, M., T. Snelder, J. Leathwick, U. Shankar, and H. Hurren, 2005. Environmental Variables for the Freshwater Environments of New Zealand River Classification. Christchurch.

Zar, J.H., 1999. Biostatistical Analysis. Prentice-Hall, New Jersey.

Proposed additional ecosystem health attributes for rivers

Adam Canning

1. For periphyton an amended attribute table is suggested. The changes shift the burden of proof for higher exceedance criteria onto Regional Council to reduce the gaming of exclusion criteria (current incentive to discharge high nutrients during months where exceedances are likely to be excluded).
    a. Should the burden of proof for the higher exceedance criteria be shifted onto Regional Council?

2. For invertebrate health, there are two attribute tables. One is based on MCI & QMCI, the other is based on Average Score Per Metric (MCI, %EPT abundance and EPT richness).
    a. Should we include both attribute tables? Or just one, if so which one?
    b. If we have the MCI & QMCI attribute table, are bottom-lines of 90/4.5 appropriate? An MCI of 90 was at the approximate inflection point with increasing nitrate based on Prof Death's nutrient paper. An MCI of 90 also represents a community that is nearly but not quite completely composed of pollution tolerant species.
    c. If we have the ASPM attribute table, which banding option should we use? Option (a) uses percentiles (75th, 50th and 25th) of current state, option (b) sets the A-band at reference condition (using logistic regression) and remaining bands equally, as per Clapcott et al., (2017). %EPT-abundance A-band calculated by converting %EPT-richness from Clapcott et al., (2017) to %EPT-abundance using regression. MFE to delete the unwanted column.

3. For fish health, the Fish-IBI has been proposed. The bands are created via percentiles (as per original IBI banding). One table includes salmonids as a positive indicator of ecosystem health given their sensitivity to water quality, reflects the RMA requirement for their protection and the Minister's cabinet paper. The other option has salmonids as a negative indicator along with other introduced species.
    a. Should trout be included as a positive or negative weighting in the Fish-IBI?

Periphyton

Value	Ecosystem health	
Freshwater Body Type	Rivers	
Attribute	Periphyton (Trophic state)	
Attribute Unit	mg chl-a/m^2 (milligrams chlorophyll-a per square metre)	
Attribute State	**Numeric Attribute State**	**Narrative Attribute State**
	Exceeded no more than 8% of samples[1,2]	
A	≤50	Rare blooms reflecting negligible nutrient enrichment and/or alteration of the natural flow regime or habitat.
B	>50 and ≤120	Occasional blooms reflecting low nutrient enrichment and/ or alteration of the natural flow regime or habitat.
C	>120 and ≤200	Periodic blooms reflecting moderate nutrient enrichment and/ or moderate alteration of the natural flow regime or habitat.
National Bottom Line	**200**	
D	>200	Regular and/or extended-duration nuisance blooms reflecting very high nutrient enrichment and/or very significant alteration of the natural flow regime or habitat.

1. May be exceeded in up to 17% of samples if shown that the exceedance would have happened at that site in natural nutrient, flow and riparian cover conditions.
2. Based on a monthly monitoring regime.

Invertebrates

Option one – MCI/QMCI:

Value	Ecosystem health		
Freshwater Body Type	Rivers		
Attribute	Macroinvertebrate Community Index and Quantitative Macroinvertebrate Community Index (Ecosystem Health)		
Attribute Unit	QMCI and MCI scores		
Attribute State	**Numeric Attribute States**		**Narrative Attribute State**
	QMCI	**MCI**	**Description**
A	≥6.5	≥130	Very clean water, indicative of pristine conditions with almost no organic pollution or nutrient enrichment.
B	≥5.5 & <6.5	≥110 & <130	Probable mild organic pollution or nutrient enrichment. Largely composed of taxa sensitive to organic pollution/nutrient enrichment.
C	≥4.5 & <5.5	≥90 & <110	Probable moderate organic pollution or nutrient enrichment. There is a mix of taxa sensitive and insensitive to organic pollution/nutrient enrichment.
National Bottom Line	4.5	90	
D	<4.5	<90	Probable severe organic pollution or nutrient enrichment. Communities are largely composed of taxa insensitive to inorganic pollution/nutrient enrichment.

1. Applies only to wadeable streams and rivers.
2. Stark JD, Maxted, JR 2007. A user guide for the Macroinvertebrate Community Index. Prepared for the Ministry for the Enviroment. Cawthron Report No. 1166. 58
3. Objectives should not be set higher than is attainable in reference conditions, with guidance from: Clapcott, J. E., Goodwin, E. O., Snelder, T. H., Collier, K. J., Neale, M. W., & Greenfield, S. (2017). Finding reference: a comparison of modelling approaches for predicting macroinvertebrate community index benchmarks. New Zealand Journal of Marine and Freshwater Research, 51(1), 44-59. doi:10.1080/00288330.2016.1265994
4. MCI and QMCI to be determined using fixed counts with at least 200 individuals surveyed using at least five Surber samplers per site annually between December and March inclusive). Sites with sediment state classes 1, 5 & 11 are to use the soft-sediment sensitivity scores. Taxonomic resolution and sensitivity scores to be use is that from Table A1.1 from:

   Clapcott, J., Wagenhoff, A., Neale, M., Storey, R., Smith, B., Death, R., … Young, R. (2017). *Macroinvertebrate metrics for the National Policy Statement for Freshwater Management*. Cawthron: Nelson, New Zealand.

5. Current state is calculated as the five-year rolling average score.

Option two – ASPM:

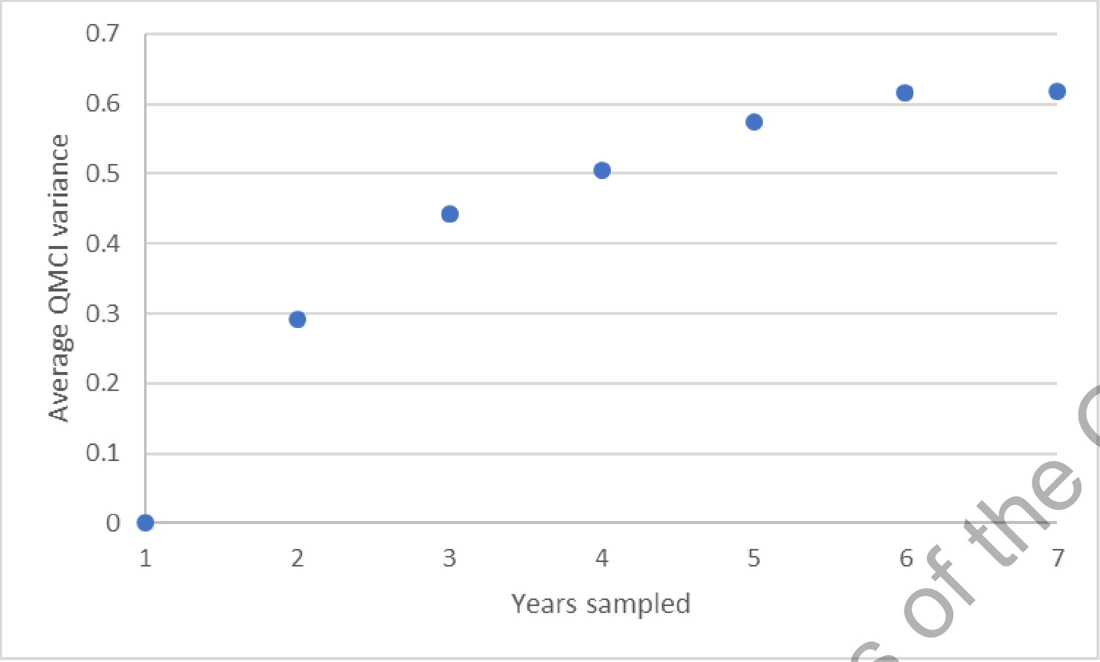
Value	Ecosystem health		
Freshwater Body Type	Rivers		
Attribute	Average Score Per Metric (Ecosystem Health)		
Attribute Unit	0-1 score		
Attribute State	**Numeric Attribute State**		**Narrative Attribute State**
	ASPM (a)	**ASPM (b)**	**Description**
A	≥0.54	≥0.6	Macroinvertebrate communities have high ecological integrity, similar to that expected in reference conditions.
B	<0.54 & ≥0.42	<0.6 & ≥0.4	Macroinvertebrate communities have mild-to-moderate loss of ecological integrity.
C	<0.42 & ≥0.27	<0.4 & ≥0.3	Macroinvertebrate communities have moderate-to-severe loss of ecological integrity.
National Bottom Line	**0.27**	**0.3**	
D	<0.27	<0.3	Macroinvertebrate communities have severe loss of ecological integrity.

1. Applies only to wadeable streams and rivers.
2. ASPM to be determined using fixed counts with at least 200 individuals surveyed using at least five Surber samplers per site annually between December and March inclusive). Sites with sediment state classes 1, 5 & 11 are to use the soft-sediment sensitivity scores. Taxonomic resolution and sensitivity scores to be use is that from Table A1.1 from:

   Clapcott, J., Wagenhoff, A., Neale, M., Storey, R., Smith, B., Death, R., … Young, R. (2017). *Macroinvertebrate metrics for the National Policy Statement for Freshwater Management*. Cawthron: Nelson, New Zealand.

3. Current state is calculated as the five-year rolling average score.
4. When normalising scores for the ASPM, use the following minimums and maximums: %EPT-abundance (0-100), EPT-richness (0-29), MCI (0-200).

Collier, K. J. (2008). Average score per metric: an alternative metric aggregation method for assessing wadeable stream health. *New Zealand Journal of Marine and Freshwater Research*, *42*(4), 367-378.



Average QMCI variance for 356 sites nationally, sampled yearly 2010-2016.

As per Stark & Phillips (2009), seasonal variability is small and annual surveys are sufficient.

As per Duggan, Scarsbrook & Quinn (2003), scores should be determined using fixed counts with at least 200 individuals surveyed using at least five Surber samplers per site annually between December and March inclusive.

As per figure above, current state should be defined as a rolling five-year average.

Fish

IBI with salmonids as positive indicator

Value	Ecosystem health	
Freshwater Body Type	Rivers	
Attribute	**Fish Index of Biotic Integrity (F-IBI)**[1]	
Attribute Unit	Score between 0-60	
Attribute State	**Numeric Attribute State**	**Narrative Attribute State**
	Average	
A	≥36	High integrity of fish community. Habitat has minimal degradation.
B	<36 and ≥28	High-moderate integrity of fish community. Habitat is mildly degraded.
C	<28 and ≥20	Moderate integrity of fish community. Habitat is moderately degraded.
National Bottom Line	**20**	
D	<20	Low integrity of fish community. Habitat highly degraded.

1. The F-IBI as defined by Joy, M. K., & Death, R. G. (2004). Application of the Index of Biotic Integrity Methodology to New Zealand Freshwater Fish Communities. *Environmental Management, 34*(3), 415-428. doi:10.1007/s00267-004-0083-0 Varied to give salmonids "honorary native" status as they are ubiquitous, are valued introduced and reflect generally good conditions, as per Joy (2015,2015 & 2013).
2. Applies only to wadeable rivers and fish are to be surveyed at least annually between December and March (inclusive) following the protocols in: Joy M, David B, and Lake M. 2013. *New Zealand Freshwater Fish Sampling Protocols (Part 1): Wadeable rivers and streams.* Palmerston North, New Zealand: Massey University.

IBI with salmonids as negative indicator

Value	Ecosystem health	
Freshwater Body Type	Rivers	
Attribute	**Fish Index of Biotic Integrity (F-IBI)**[1]	
Attribute Unit	Score between 0-60	
Attribute State	**Numeric Attribute State**	**Narrative Attribute State**
	Average	
A	≥34	High integrity of fish community. Habitat has minimal degradation.
B	<34 and ≥26	High-moderate integrity of fish community. Habitat is mildly degraded.
C	<26 and ≥16	Moderate integrity of fish community. Habitat is moderately degraded.
National Bottom Line	**16**	
D	<16	Low integrity of fish community. Habitat highly degraded.

1. The F-IBI as defined by Joy, M. K., & Death, R. G. (2004). Application of the Index of Biotic Integrity Methodology to New Zealand Freshwater Fish Communities. *Environmental Management, 34*(3), 415-428. doi:10.1007/s00267-004-0083-0
2. Applies only to wadeable rivers and fish are to be surveyed at least annually between December and March (inclusive) following the protocols in:
   Joy M, David B, and Lake M. 2013. *New Zealand Freshwater Fish Sampling Protocols (Part 1): Wadeable rivers and streams.* Palmerston North, New Zealand: Massey University.

IBI scores are statistically different with and without trout at a positive health indicator.

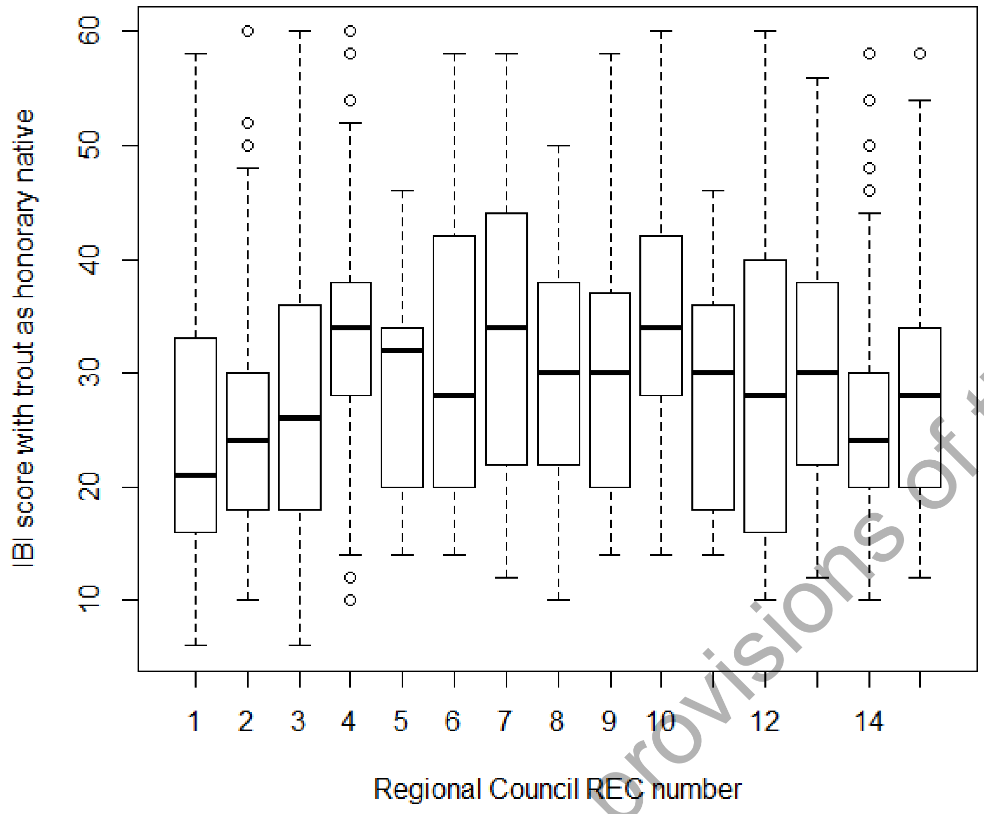
Anova: Single Factor

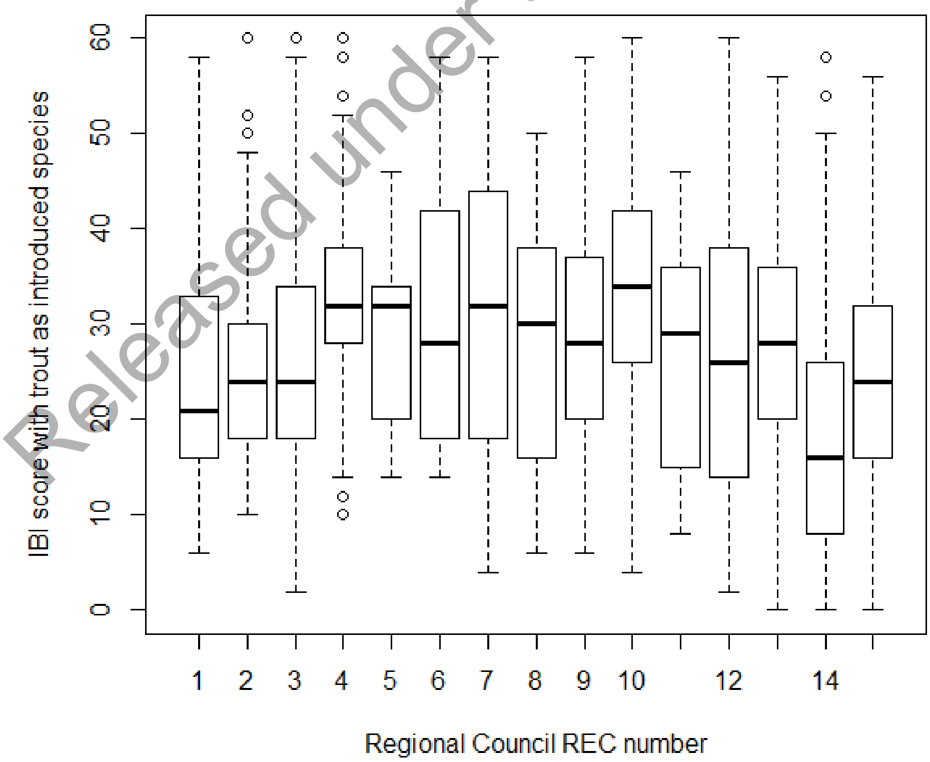
SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Column 1	3009	87130	28.95646	121.8867
Column 2	3009	79126	26.29644	161.319

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	10645.4	1	10645.4	75.17785	5.46E-18	3.843005
Within Groups	851882.9	6016	141.6029			
Total	862528.3	6017				


IBI score with trout as honorary native
10
20
30
40
50
60
1
2
3
4
5
6
7
8
9
10
12
14
Regional Council REC number


IBI score with trout as introduced species
0
10
20
30
40
50
60
1
2
3
4
5
6
7
8
9
10
12
14
Regional Council REC number

Ecosystem processes

Value	Ecosystem health				
Freshwater	Rivers				
Attribute	Ecosystem metabolism				
Attribute Unit	g O_2 m^{-2} d^{-1} (grams of dissolved oxygen per square metre per day)				
Attribute State	**Numeric Attribute State**				**Narrative Attribute State**
	Gross primary production		Ecosystem respiration		
	Non-wadeable	Wadeable	Non-wadeable	Wadeable	
A	≤3.0	≤3.5	1.6-3.0	1.6-5.8	No evidence of an impact on ecosystem metabolism.
B	>3.0 and <5.5	>3.5 and <5.0	>1.0 and <1.6 Or >3.0 and <8	>1.2 and <1.6 Or >5.8 and <7	Mild effect on ecosystem metabolism.
C	≥5.5 and ≤8.0	≥5.0 and ≤7.0	≥0.6 and ≤1.0 Or ≥8.0 and ≤13.0	≥0.8 and ≤1.2 Or ≥7.0 and ≤9.5	Moderate effect on ecosystem metabolism.
National Bottom Line	**8.0**	**7.0**	**≥0.6 or ≤13.0**	**≥0.8 or ≤9.5**	
D	<8.0	<7.0	<0.6 or >13.0	<0.8 or >9.5	Severely impaired ecosystem metabolism.

1. Derived from 7 consecutive days of continuous dissolved oxygen monitoring, though objective applies year-round.
2. Young, R. G., Matthaei, C. D., & Townsend, C. R. (2008). Organic matter breakdown and ecosystem metabolism: functional indicators for assessing river ecosystem health. Journal of the North American Benthological Society, 27(3), 605-625. doi:10.1899/07-121.1
3. Clapcott JE 2015. Development of management bands for ecosystem metabolism in nonwadeable rivers. Prepared for Waikato Regional Council. Cawthron Report No. 2770. 21 p. plus appendix.