Australian Rivers Institute
Nathan campus, Griffith University
170 Kessels Road
Nathan, Queensland 4111
Australia

Telephone: [redacted]

www.rivers.edu.au

Gold Coast Logan Mt Gravatt Nathan South E

Dr Jen Price
Senior Analyst, Freshwater Science & Technology
Ministry for the Environment – Manatū Mō Te Taiao
23 Kate Sheppard Place, PO Box 10362, Wellington 6143

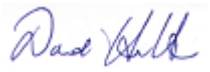
Dear Jen,

Thank you for forwarding documents to me that enabled me to conduct a review related to the management of nutrients in rivers in the NPSFM. I interpret that you overarching questions were:

- Does the evidence support the conclusions of the advisory group?
- Are the options defensible and is the evidence base for both sufficient and defensible?

I have provided further details of the scope on the following page, following by my interpretation.

Regards,



David Hamilton

Review of management of nutrients in rivers as part of the NPSFM – Scope (Dr Jennifer Blair)

Abstracted material from the following email:
From: Jennifer Price
Sent: Monday, 20 May 2019 9:56 AM
To: David Hamilton (david.p.hamilton@griffith.edu.au) <david.p.hamilton@griffith.edu.au>
Cc: ken.taylor@agresearch.co.nz
Subject: FW: Email for review please: Peer review of nutrient recommendations: context and key papers

Context

The Science and Technical Advisory Group has discussed the management of nutrients in rivers in the NPSFM, in particular in relation to a proposal from Russell Death and Adam Canning to introduce attributes for nitrogen and phosphorus that have been derived based on a weight-of-evidence approach.

Ken has summarised the proceedings of the STAG meetings in relation to nutrients as follows:

- There was very strong support for the need to set nutrient thresholds to manage nutrient health
- There was broad agreement that Russell's methodology was robust
- There was strong agreement that further work needed to be done to answer a number of technical questions, including whether or not whether nutrient bottom lines should be spatially differentiated.

These matters were discussed further at the March and April STAG meetings. At one of those meetings it was clear that there was strong support for having a single set of numbers for nitrogen and phosphorus, and also strong agreement (but not unanimity) for Russell's numbers. Through those discussions and by email exchanges a number of complexities have become apparent. These concern the forms of nitrogen and phosphorus that should be prescribed (e.g., total v dissolved, NO3-N v DIN), aspects of the application of the weight-of-evidence methodology (including relative weightings of the different strands of evidence), and the content of attribute narratives. It was also clear that there were some STAG members who were of the view that there were unresolved questions about the weight-of-evidence approach, although Ken's assessment is that nearly everyone in the group is comfortable with the general are in which the numbers have landed.

Key documents to review

File name	What this is
STAG 26 Feb Death et al 2018 Nutrient criteria for NZ rivers	Paper detailing proposed nutrient criteria – forms basis for further STAG discussions
Adam's new analysis	Email containing the results of modelling analyses conducted by Ton Snelder – showing little relationship between N and MCI
RE: Nutrient sub-group – decision needed about sub-group summary	Email containing the results of modelling analyses conducted by Adam Canning – showing N is a key predictor of MCI
STAG 16 April agenda and priority papers	See paper detailing analysis of spatial variation in nutrient-MCI relationship and maps showing where proposed tables would apply, agenda item 6, pg 30

	Also contains agenda, draft minutes of previous meeting and key papers.
Your query about nutrient narrative sub-group correspondence	Summary of proposal: A/B band boundary for nitrate toxicity attribute becomes new bottom line
Proposed nutrient attributes 6 May JWB ver (2)	STAG's proposed nutrient attribute tables
TO COME: Methods and results of analysis to produce nutrient attribute tables for ecosystem health	
TO COME: Summary of STAG recommendations	

Questions for review

Does the evidence support the conclusions of the advisory group?

Are the options defensible and is the evidence base for both sufficient and defensible?

Options

- Status quo (no new nutrient tables for ecosystem health): periphyton attribute with toxicity attributes
- Weight of evidence approach and resulting numbers for DIN and DRP
- Single table based on nitrate toxicity criterion that gives a high level of protection (nitrate toxicity A/B band becomes bottom line)

Constraints on review - STAG has recommended a single table, and the option of designing attribute tables that vary spatially is not feasible under the time frames we have.

Review of management of nutrients in rivers as part of the NPSFM

Provision for additional attributes

It is generally accepted that nutrient thresholds are required to manage eutrophication and health of New Zealand rivers and streams according to the National Policy Statement for Freshwater Management, specifically in relation to Objective B1 of the NPSFM:

> To safeguard the life-supporting capacity, ecosystem processes and indigenous species including their associated ecosystems of fresh water, in sustainably managing the taking, using, damming, or diverting of fresh water.

Nutrient thresholds would add to the existing attribute of periphyton for assessing ecosystem health. It is widely recognised that periphyton is a useful attribute for stream health assessment but it is not applicable to soft-bottomed streams and is also influenced by several other processes besides nutrient concentrations, which can make interpretations of stream health difficult and subject to variability. Nutrients, as a fundamental driver of eutrophication, may potentially better reflect opportunities to better manage stream health through, for example, more direct connectivity of in-stream nutrient concentrations to land use management. I am interested to know if toxicity thresholds would become irrelevant with the introduction of nutrient thresholds for managing eutrophication for ecosystem health, given that these thresholds would almost certainly be considerably lower than the toxicity values and that, in my opinion, nutrient toxicity thresholds have been confusing and sometimes misinterpreted:

Recommendation 1: Clarification could be sought that if nutrient attributes are introduced to manage eutrophication for ecosystem health purposes, then nutrient toxicity attributes would no longer be required.

Assessment of nutrient species

It can be difficult to interpret how nutrients drive productivity in aquatic systems because of the multiple fluxes that influence concentrations. Total nitrogen and total phosphorus concentrations tend to be easier to interpret for this purpose because they act somewhat more conservatively. Both total nitrogen and total phosphorus have a limited number of processes affecting their concentrations; they can sediment out or be resuspended, and for nitrogen there may be interactions with gaseous phases (e.g., via denitrification or nitrogen fixation). However, interpretation of concentrations of bioavailable nutrients, defined operationally by nitrate-N, ammonium-N and dissolved reactive phosphorus, are far more complex because of a multitude of different processes (e.g., desorption from particulate phases, nitrification, denitrification, nitrogen-fixation, sedimentation, etc.) affecting their concentrations.

The NPSFM uses total nitrogen and total phosphorus to assess ecosystem health of lakes because these constituents are generally strongly correlated with phytoplankton chlorophyll *a*. Indeed, phytoplankton quotas of nitrogen and phosphorus often constitute a large fraction of the total nitrogen and phosphorus in water. By contrast, dissolved nutrient concentrations may not be positively correlated with chlorophyll *a* and, in work I have been involved with (Wood et al. 2014) can even be strongly depleted in the presence of high chlorophyll *a,* driven by demand from high primary producer biomass and with sufficient residence time for quasi-equilibration of biomass to the inorganic nutrient supply. Even in estuaries, it has been shown that phytoplankton biomass may vary inversely with dissolved inorganic nutrient concentrations (Chan et al. 2001).

If dissolved inorganic nutrient concentrations are considered to be a reliable indicator of stream health impacts related to eutrophication because, then the premise is presumably that mass fluxes of these nutrients generated by runoff and transport exceed the rate of in-stream processing (e.g., from periphyton uptake) that might otherwise reduce concentrations to low levels. In addition, use of dissolved nutrients may obviate some of the problems associated with sometimes large components of total nutrients being largely biologically unavailable, for example, in dissolved organic form (e.g., DON).

Recommendation 2: There should be clear justification for using dissolved inorganic nutrients versus total nutrients as an attribute. Assessments using dissolved inorganic nutrients may need to consider the temporal and spatial variability of dissolved inorganic nutrients, for example associated with stream discharge.

In much of the information that has been presented to me, two nutrient species have been examined as attributes in the NPSFM: nitrate and dissolved reactive phosphorus. However, ammonium is a bioavailable form of dissolved inorganic nitrogen that is also used as a nutrient by aquatic primary producers. An assumption has been made in some of the reports presented to me that ammonium may be neglected because stream waters are generally well oxidised. However, if there are discharges high in ammonium, slow-flowing waters subject to high organic loads, or in-stream (sediment) releases, then ammonium concentrations may be elevated. Ammonium is also considered to be preferentially taken up over nitrate by aquatic primary producers. I consider a deeper evaluation of this issue may be required by the STAG.

Recommendation 3: If dissolved inorganic nutrients are to be used in the attribute table, then consideration should be given to use of concentrations of dissolved inorganic nitrogen (DIN = NO_3-N + NH_4-N) as a nitrogen attribute in preference to NO_3-N, to reflect the nitrogen supply available to aquatic primary producers.

Evaluation of Death et al. 'Clean but not green: a weight-of-evidence approach for setting nutrient criteria in New Zealand rivers' as a basis for nutrient attribute criteria

The report by Death et al. provides a weight-of-evidence approach for selecting threshold concentrations of nitrate and dissolved reactive phosphorus designed to correspond to attribute states used in the NPS-FM. The approach provides a basis for relating different stream health proxies to concentrations of these two nutrient species. My opinion is that weight-of-evidence approach to stream health assessment is valid but that in this instance the methodology is strongly reliant on repeated use (8-9x) of slightly different assessment methods for the same health indicator group (macroinvertebrates). Compounding the limited biota assessment is an additional weighting factor (2x) to further increase the influence of macroinvertebrates in the weight of evidence, even when sometimes (case 5, MCI) they explained as little as 8% of the variation in NO_3-N concentrations. The report 'Nitrate-nitrogen and Dissolved Reactive Phosphorus (Rivers and groundwater)' presented to the Science and Technical Advisory Group meeting on 26 February 2019 by A. Canning provides an explanation that weighted averaging was based on whether linkages between nutrients and an ecosystem health metric were direct (2x) or indirect (1x). I remain unconvinced by this explanation as only the periphyton chlorophyll *a* metric responds in a direct way to nutrients, not macroinvertebrates. I take a level of comfort with relationships of chlorophyll *a* to nutrients (r^2 = 0.3 for DRP and NO_3-N) but note that the IBI has very low predictive power (r^2 = 0.09 for DRP and r^2 = 0.04 for DRP) and its significance ($p < 0.0001$) is largely irrelevant because of the very large sample size.

Recommendation 4: Consideration could be given whether it is valid to use different metrics of the same (or similar) indicator group to build weight-of-evidence for nutrient concentration thresholds.

Recommendation 5: Consideration could be given to setting an acceptable (statistical) cut off for including an indicator to provide weight-of-evidence for establishing nutrient concentration thresholds. Levels of significance (p values) are irrelevant for many of the large data sets used for the weight-of-evidence approach.

Evaluation of 'Proposed Nutrient Attribute Tables for the NPS-FM' (21.05.2019) as a basis for nutrient attribute criteria

The recent report that I was provided (Proposed Nutrient Attribute tables for the NPS-FM (21.05.2019)) deals with only two macroinvertebrate attributes, the QMCI/MCI and ASPM, and it proposes the use of GPP, ER and Cotton K as components of the weigh-of-evidence approach. It indicates that weighting factors are not applied to the components used in the weight-of-evidence. I consider that some of the questions raised above should also be considered with reference to this report: do defined statistical thresholds (e.g., based on R^2) allow GPP, ER and Cotton K to be included? Fig. 3 (Cotton K vs Nitrogen), for example, looks more like a random distribution to me, with the nitrogen variable apparently transformed, as its scale shown in Fig. 3 is incorrect. I was also not strongly convinced by the quantile regressions between Fish IBI and nitrate-nitrogen, and why 90^{th} and 80^{th} percentiles were chosen to correspond to particular attribute states (A/B and C/D).

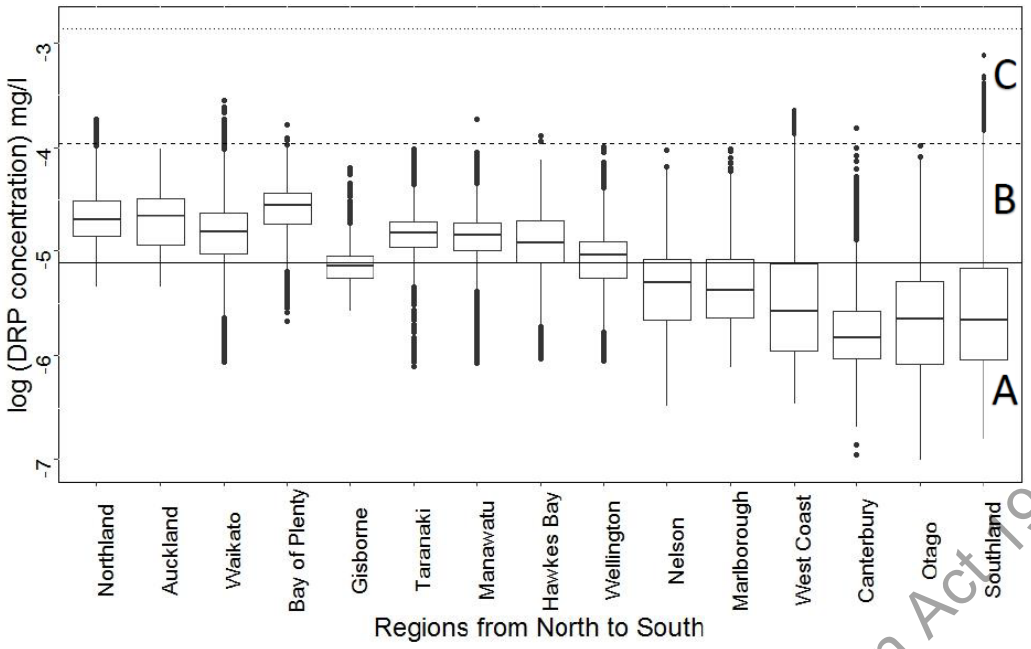
Comparison of *An exploration of factors predicting macroinvertebrate indices* (A. Canning) and results contained in an attachment to an email with subject line '*Adam's new analysis*' (T. Snelder)

I do not feel that I can adequately comment on the relative merits of these two analyses (by A. Canning and T. Snelder), which present quite different statistical outcomes. As mentioned by Snelder, results from Boosted Regression Tree analysis and Random Forest Models should be similar. Such conflicts in an apparently quantitative analysis are of concern and need to be resolved for the credibility of the advice being provided to the STAG. A logical first step is to make sure there is commonality and agreement on the data sets being used for the analysis.

Recommendation 6: Alignment of data sets used by scientists needs to be carried out urgently so that there is greater consistency of statistical information and analysis provided to the STAG. At the very least, a common example data set should be used to show that different statistical analyses are broadly in agreement.

General comments

I was a little surprised that more consideration was not given to the analyses by McDowell et al. (2013). In some cases there even seemed to be misinterpretation of some of their results (e.g., "…there was not substantial variation in natural nutrient concentrations between regions" (A. Canning), which, for example, is at odds with McDowell et al. finding that DRP reference concentrations for the volcanic acidic lake category were strongly elevated, although they were not likely to exceed the bottom line proposed from the weight-of-evidence approach (taking this value to be 0.038 mg L^{-1}). I am aware, however, of some volcanic acidic streams that may be considered to be in a reference state, and would exceed this value. In addition, it is fairly clear that there are consistent differences between North and South Islands:



Recommendation 7: Work by McDowell et al. (2013) and recently by Abell et al. (also with McDowell) should be re-examined by the STAG for the purpose of deriving spatial variations in stream reference nutrient concentrations across New Zealand.

Such an analysis could also include longitudinal gradients within streams. The work may be able to be accomplished quickly based on prior analyses by T. Snelder. It could be used to delineate proportions of sites falling into one of the proposed weight-of-evidence nutrient classification attribute tables.

Conclusions

My opinion is that the numerical values of nutrient concentrations derived for supporting thresholds for different Attribute States (i.e., A/B, B/C and C/D) given in the Death et al. (and/or Canning) reports match reasonably well with my own interpretation of whereabouts the concentrations would 'fall out'. There is strong evidence for additional attributes besides periphyton and nutrient toxicity to manage stream ecosystem health. However, there needs to be a much stronger evidence base:

- For dissolved versus total nutrient concentrations (including DIN vs NO_3-N)
- For inclusion of different proxies as variables in the weight-of-evidence approach (also a quantitative statistical justification)
- For agreed common data sets that should allow different statistical approaches to be compared directly, and agreed upon
- To include referral and cross-checking with reference concentrations so that spatial variations are accounted for in the need to develop consistent national attribute tables.

References

McDowell, R. W., Snelder, T. H., Cox, N., Booker, D. J. & Wilcock, R. J. 2013. Establishment of reference or baseline conditions for chemical indicators in New Zealand streams and rivers relative to present conditions. *Marine and Freshwater Research* 64: 387-400. (See also similar content in an AgResearch report by McDowell, R. W., Snelder, T. H., Cox (2013) provided to the Ministry for the Environment).