Environment Southland - Review of environmental models

To	Environment Southland c/o Wilma Falconer and Karen Wilson
Cc	Ned Norton
From	Hamish Brown, Tim Davie, Andrew Fenemor, Bethanna Jackson, Richard Muirhead, Mike Scarsbrook, Andrew Schollum, Ken Taylor (Chair).
Date	17 October 2022
Topic	**Science Review Panel memo to Environment Southland**

Purpose

1. We have been asked to conduct an independent review of the modelling process followed by Environment Southland to estimate the baseline state of water quality in Murihiku/Southland and gain a general understanding of the reduction in contaminant levels required to achieve draft freshwater objectives identified for the region. The scope of this review includes consideration of the robustness and credibility of the contaminant modelling relating to nitrogen, phosphorus, and suspended sediment/visual clarity in surface waters, and assessment of the appropriateness of the modelling process for informing policy development.
2. Decisions have yet to be taken on how model outputs will be used in the regulatory plan-making process. This and a range of other matters have been explicitly placed out of scope for our review, including how draft freshwater objectives are determined, what risk thresholds/exceedance criteria are used to determine limits and where contaminant limits are set, and whether the achievability of objectives is factored into the setting of limits. Nevertheless, Environment Southland has invited us to provide commentary should our review identify matters that could have a bearing on regulatory design or that we believe should be considered by decision makers as they consider translating model outputs into policy.

Method

3. In undertaking this review, we:
    a. Participated in a virtual pre-review background/briefing session with Environment Southland Staff.
    b. Reviewed the following reports:
        i) Snelder, T., (2021) *Assessment of Nutrient Load Reductions to Achieve Freshwater Objectives in the Rivers, Lakes and Estuaries of Southland Including Uncertainties: To inform the Southland Regional Forum process*, Land Water People
        ii) Plew, D., (2020) *Models for evaluating impacts of nutrient and sediment loads to Southland Estuaries: To inform the Southland Regional Forum*

process, National Institute of Water & Atmospheric Research, NIWA CLIENT REPORT No: 2020216CH

iii) Neverman A., Smith H., Herzig A., Basher L., (2021) *Modelling baseline suspended sediment loads and load reductions required to achieve Draft Freshwater Objectives for Southland*, Manaaki Whenua – Landcare Research, Contract Report: LC3749

c. Read key papers, referenced in the reports described in paragraph 3b, that described previous work undertaken as a precursor to the modelling processes e.g., the processes used to determine the attributes and bands set out in the 'National Objectives Framework.'

d. Reviewed a series of reports/memos provided by the modelling team at Environment Southland that:

    i. Described the overall aims of Environment Southland's work programme,

    ii. Outlined the overall approach Environment Southland has taken to identify freshwater objectives and estimate the magnitude of change required to meet those objectives, and

    iii. Summarised the key conclusions and implications arising from the modelling work and identified questions arising from Environment Southland's own review of the modelling undertaken and comments received from key stakeholders.

e. Reviewed commentary on the modelling programme provided by key stakeholders.[1]

f. Participated in two day-long review panel workshops in-person in Christchurch and three virtual workshops/question-answer sessions with the authors of each of the modelling reports.

g. Engaged in email dialogue between members – sharing perspectives, raising questions, and debating points of detail – and reviewed and provided written and verbal comments on drafts of this memo.

4. At our in-person workshops we analysed each model/modelling project, following the order set out in paragraph 3b, using a series of 'prompts' (see **Attachment 1**) to help guide[2] the assessment, coordinate the feedback of panel members, and to structure discussion on each of the models/modelling projects.

5. We began each session by assessing whether the modelling undertaken was appropriate in this context, and whether it was undertaken in accordance with general principles of best practice for modelling. We then spent the remainder of each session assessing specific matters of technical, methodological, and scientific detail, considering the following questions: whether the modelling is conceptually, technically, and scientifically robust, at what scale the models can be used with confidence, and whether the modelling is appropriate for the purpose Environment Southland intends to use it.

6. This memo provides an overview of our analysis and presents the conclusions of our review.

[1] The panel would like to give particular thanks to James Cook University, which provided helpful commentary on behalf of Fish and Game New Zealand, and Dairy NZ for its comprehensive and thought-provoking analysis of the model reports and underpinning data.

[2] Some questions/prompts were less relevant for some reports, and we used them as a flexible rather than prescriptive guide.

Conclusions

1. The models used by Environment Southland to estimate contaminant loads are, in our opinion, conceptually sound – the choice of models and the modelling architecture are appropriate given the purpose for which the models are being used at this stage in Environment Southland's programme for implementing the National Policy Statement for Freshwater Management.
2. We are satisfied the modelling undertaken by Environmental Southland provides a suitable foundation for estimating the reduction of contaminant levels required to achieve draft freshwater objectives set by Southland Regional Council and the board of Te Ao Marama Incorporated in December 2020. Specifically, we consider the models used by Environment Southland provide an appropriate basis for:
    - making quantitative estimates of current contaminant loads,
    - extrapolating in space to predict conditions around the region and infer compliance with objectives set at a broad geographic scale, rather than just at monitoring locations,
    - identifying critical points of sensitivity and key contaminant sources within catchments,
    - establishing where more targeted observation, monitoring, and modelling are required to build a more robust understanding of environmental state, trends, pressures, and responses prior to setting contaminant limits.
3. We note that periphyton growth is highly spatially variable and there is significant uncertainty in the model estimates regarding predicted compliance with draft freshwater objectives. Predicted levels of compliance are highly sensitive to the level at which exceedance criteria are set. Exactly where exceedance criteria are set is a choice that sits with the council and local community – neither the modellers nor this panel have a view on where these criteria should be set, but we are confident that the model outputs provide a robust basis for informing policy decisions.
4. The methods used by the modelling teams were replicable, and data sources and assumptions were described or cited in sufficient detail to facilitate interrogation. We consider the approach to modelling has been designed appropriately to allow Māori and non-Māori knowledge systems to grow together, work alongside each other, coordinating when it makes sense to do so and standing apart but alongside each other when that is appropriate.
5. The modelling undertaken by Environment Southland is consistent with the requirements set out in the National Policy Statement for Freshwater Management and related guidance documents. We note, and consider it appropriate, that key modelling steps, particularly those associated with classifying water quality in relation to National Objectives Framework bands, relied on algorithms adopted by the Ministry for the Environment when developing the National Policy Statement for Freshwater Management.
6. We emphasise that where output from one model serves as input for another, uncertainty may propagate, increasing the uncertainty of linked model predictions. In these instances, decision-makers will need to be particularly cautious and avoid using the model outputs beyond their intended scope. Ideally Environment Southland's model reports would contain a more comprehensive explanation of the sources and nature of uncertainty in the modelling and spell out the implications of uncertainty for model users and for the potential scope of model application. We note that many of the modelling steps taken require the modellers to make technical/scientific judgements. It is important that the decision makers using these "linked" model outputs are aware of the sensitivity of model predictions to these judgements.

7. We note that Environment Southland, and all other regional councils in New Zealand, will need to continue to strive to integrate data from sampling, observation, and modelling to provide the best available information to decision-makers responsible for implementing the National Policy Statement for Freshwater Management. This implies that Environment Southland, and all other regional councils in New Zealand, will need to commit to continually improving the information base upon which they make decisions – this will require significant resourcing and continued support from central government.

Is the modelling appropriate in this context?

Why use models?

7. Many factors affect the health of freshwater bodies, and these factors interact in complex ways – the same mix of contaminants or environmental pressures at different times and in different places can have different effects on the health of a system. To gain an understanding of the processes that influence freshwater ecosystems, determine the current state, identify trends and pressures, and predict future conditions, regional council environmental managers must draw on and integrate data from environmental monitoring, field observations, and environmental models.
8. These sources of information are interdependent. It is often necessary, for instance, to use statistical models to enable calculation of water quality indices from monthly field samples, or to use models to explain the results of sampling programmes,[3] and it is necessary to have monitoring data to build, train, and deploy environmental models.
9. Environmental models play an important role in achieving Te Mana o te Wai[4] as they allow managers to generate useful insights into the interactions between complex natural and human systems, identify system drivers (causes), and forecast future conditions (outcomes) under a range of different management scenarios, and at a range of spatial and temporal scales.
10. Data from Environment Southland's environmental monitoring programme indicate there are widespread water quality issues in the region and that the health of many freshwater bodies does not meet community expectations as reflected in draft water quality objectives for the region. Although Environment Southland's environmental monitoring programme provides useful information on the current state of the environment, the limited spatial and temporal coverage of monitoring data across the region restricts the ability of managers to gain an integrated region-wide picture of issues and options. Nor can Environment Southland's monitoring data, on its own, enable policy makers to assess the likely effectiveness of different management options and available mitigation packages (i.e., predict future state under different management settings and determine what actions are necessary to bring a waterbody into compliance with limits specified in plans.)

[3] For example, the need to quantify land use intensity to explain increases in monitored nutrient concentrations and calculated loads.

[4] The National Policy Statement for Freshwater Management (2020) establishes Te Mana o te Wai as the fundamental concept driving freshwater management in New Zealand. Te Mana o te Wai requires regional councils to prioritise, first the health and well-being of water bodies and freshwater ecosystems, second the health needs of people, and third the ability for people and communities to provide for social, economic, and cultural well-being.

11. Given financial and practical constraints, and statutory deadlines for amending plans, it is not possible for Environment Southland to sufficiently extend the number of monitoring sites and frequency of sampling to generate the spatial and temporal accuracy of information required to achieve desired freshwater objectives with confidence. In any case, data from this monitoring will often need to be passed through models to make sense of the results and generate meaningful insight into environmental state, trends, and likely future states under different objectives and management scenarios.

12. For these reasons we are satisfied it is appropriate for Environmental Southland to use available monitoring data to develop, train, and validate environmental models with the objective of estimating the regional and catchment-scale reductions of contaminant levels likely to be required to achieve draft freshwater objectives set by the community.

Is the approach to modelling conceptually sound?

13. This stage of Environment Southland's modelling programme has been designed to give an understanding of the difference between target attribute states ("draft freshwater objectives"), the current state of the environment (current attribute state and contaminant loads), and the reductions in contaminant loads predicted to be needed to achieve the draft target attribute states – in other words, the 'size of the gap.'

14. By defining the 'size of the gap', Environment Southland aims to:

    a. Gain a general understanding of the reduction in contaminant levels required to meet freshwater objectives – characterising the scale of the issue facing the region,

    b. Establish whether and where there are differences across the region in terms of the scale of the issue and the contaminants driving water quality pressures, and

    c. identify where more detailed modelling and targeted monitoring is required to better understand local "hotspots" and support accurate, effective management responses.

15. Consistent with the requirements of the National Policy Statement for Freshwater Management 2020 (NPSFM), Environment Southland is considering a broad range of contaminants that affect water quality. This involves undertaking investigations into specific contaminants and contaminant pathways and focussing on areas of known sensitivity based on an understanding of the context within each of the region's catchments and freshwater management units. We note that, while microbial contamination (*E. coli*), groundwater contamination and nitrate contamination are under investigation, this work was not within scope of our review, which focussed on Environment Southland's investigation into nitrogen, phosphorus, and suspended sediment/visual clarity in surface waters.

16. Environment Southland has followed a spatially integrated approach to define the 'size of the gap' for nitrogen, phosphorus, and suspended sediment/visual clarity. This provides managers with a picture of nutrient and sediment issues as they relate to multiple attributes and ecosystem-types distributed across the entire region. It would not be possible to generate such a geographically extensive and integrated picture of water quality if the council were restricted to using site-specific monitoring data alone. To undertake this task Environment Southland fed its monitoring data into a suite of linked models (see **Figures 1 and 2 on pages 8 and 9** for schematic representations of how the models link together):

a. The modelling described in the Snelder (2021) report estimates nitrogen and phosphorus load reductions required to achieve objectives for rivers (periphyton and nitrate toxicity), lakes (total nitrogen, total phosphorus and hence by assumption lake phytoplankton), and estuaries (macroalgae and/or estuary phytoplankton). This model draws on output from the Abell et al (2020)[5] component model to estimate lake concentration based on input loads calculated by Snelder, and output from the Plew (2020) estuary component models to derive loads for total nitrogen and total phosphorus necessary to achieve objectives for macroalgae and/or phytoplankton in estuaries.

b. The modelling described in the Neverman et al., (2021) report estimates suspended sediment load reductions required to achieve visual clarity objectives in rivers. This model draws on output from a 'Random Forest' regression model supplied by Snelder to establish baseline visual clarity, and uses the Hicks et al., (2019)[6] equation to derive the per cent change in suspended sediment load needed to achieve the target attribute state for visual clarity in rivers.

c. The outputs from Neverman et al., (2021) are used by the modelling described in the Plew (2020) report to identify relative differences between baseline and "natural cover" sedimentation rates in estuaries, and to coarsely test whether the sediment load reductions to achieve visual clarity in rivers could also be sufficient to achieve draft sedimentation rate objectives in estuaries.

17. In this memo we provide general commentary on the conceptual and scientific basis of the models described in the three model reports that are within scope of this review. In general terms, we consider that the choice of models and the modelling architecture – the way models are linked – is appropriate given:

a. the purpose for which the models are being used at this stage in Environment Southland's programme – characterising the 'size of the gap' between current and target states, and

b. the need to maintain consistency between the approach underpinning the NPSFM and the approach to implementing its policies (we note that some of the choices integral to the design of the NPSFM, for instance the decision to manage sediment via compliance with visual clarity target attribute states, rely on previously published and peer-reviewed assumptions and interrelationships that have inherent uncertainty associated with them.)

18. Similarly, we are satisfied the modelling undertaken by Environmental Southland to date provides a conceptually sound basis for:

a. estimating contaminant loads and extrapolating spatially to:

i. predict conditions around the region, and

ii. infer compliance with objectives set at a broad geographic scale (e.g., no more than 20 per cent of sites within a freshwater management unit or

[5] Abell, J.M., P. van Dam-Bates, D. Özkundakci, and D.P. Hamilton, 2020. Reference and Current Trophic Level Index of New Zealand Lakes: Benchmarks to Inform Lake Management and Assessment. New Zealand Journal of Marine and Freshwater Research:1–22.

[6] Hicks M, Haddadchi A, Whitehead A, Shankar U, 2019. Sediment load reductions to meet suspended and deposited sediment thresholds. NIWA Client Report 2019100CH prepared for Ministry for the Environment.

reporting catchment experience periphyton growth that exceeds a given threshold.)

b. identifying critical points of sensitivity and key sources of contaminants within catchments, and establishing where more targeted observation, monitoring, and modelling are required to build a more robust understanding of environmental state, trends, pressures, and responses.

19. There are, however, many instances where data are sparse and in our specific comments on each individual model report, we note areas where there is a high degree of uncertainty surrounding model predictions. In certain instances, where output from one model serves as input for another, uncertainty may propagate, increasing the uncertainty associated with model predictions. In these instances, decision-makers will need to be particularly cautious to avoid using the data generated by these models beyond their intended scope.

20. We note that many of the individual steps and/or models used in this analysis require modellers to make technical/scientific judgements when developing algorithms (e.g., erosivity and vegetation cover coefficients.) We also note these models have been developed within a policy framework that requires policymakers to make normative judgements. These judgements are key input parameters for the models and have a significant influence on model predictions regarding the level of contaminant reductions required to meet community objectives (e.g., the percentage threshold chosen when determining spatial exceedance criteria.)

21. We consider that these technical/scientific judgements have been made in an informed manner, are sufficiently transparent, and that the individual algorithms included within these models are defensible and represent the best available information at the time. We do note, however, that it is important decision makers using these “linked” model outputs are aware of the impact these judgements have on model calculations and outputs.

22. After much debate, we have concluded that it is difficult to see what more Environment Southland could do in the short term to manage uncertainty when balancing data and resource limitations against the requirements of the NPSFM.

23. In this regard, we note that Environment Southland, like all other regional councils in New Zealand, will need to continue to strive to integrate data from sampling, monitoring, and modelling programmes to provide the best available information for decision-making, as required by the NPSFM. This implies that Environment Southland, and all other regional councils in New Zealand, will need to commit to continually improving the information base upon which they make decisions. This will require significant resourcing, and continued support from central government.

Figure 1: Nutrient load reductions assessment components

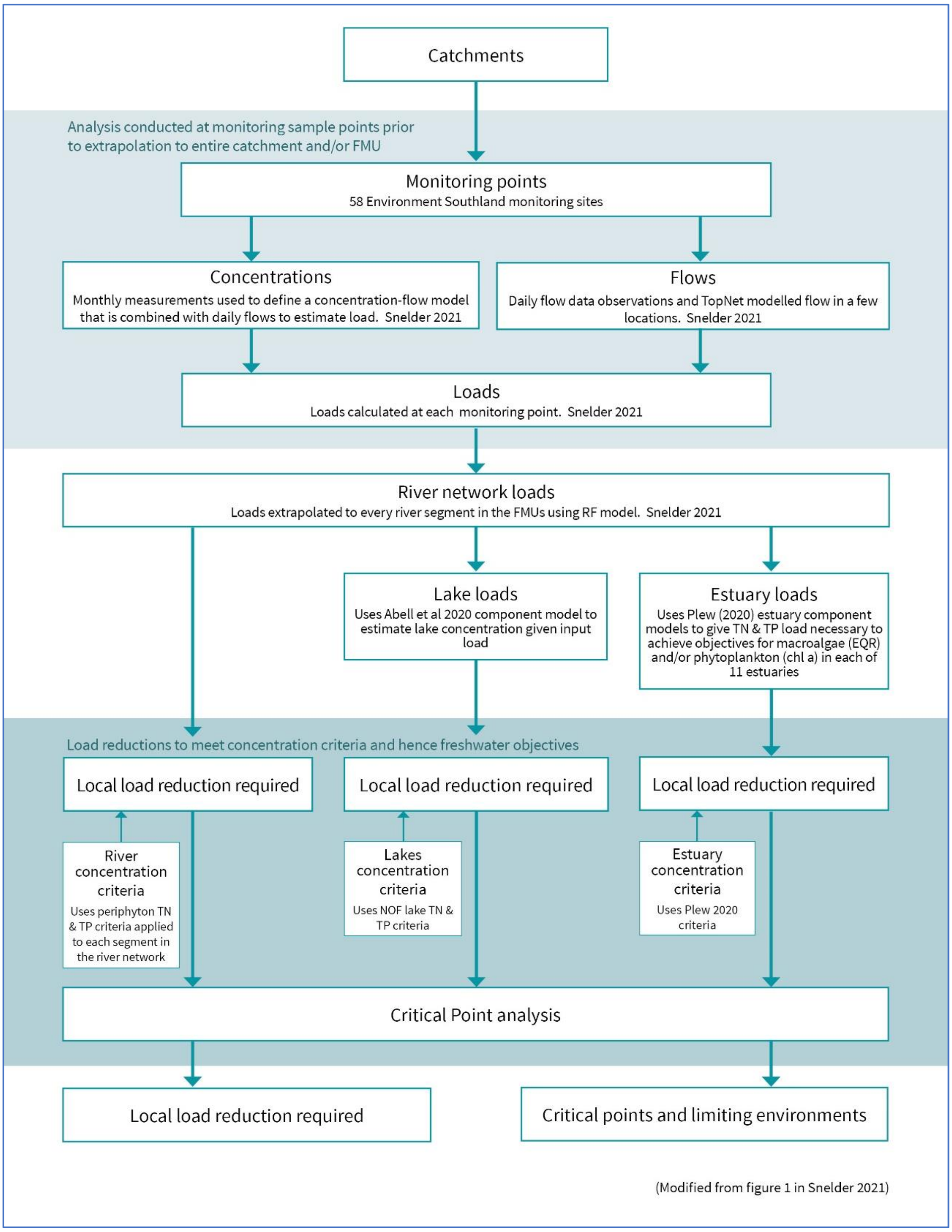
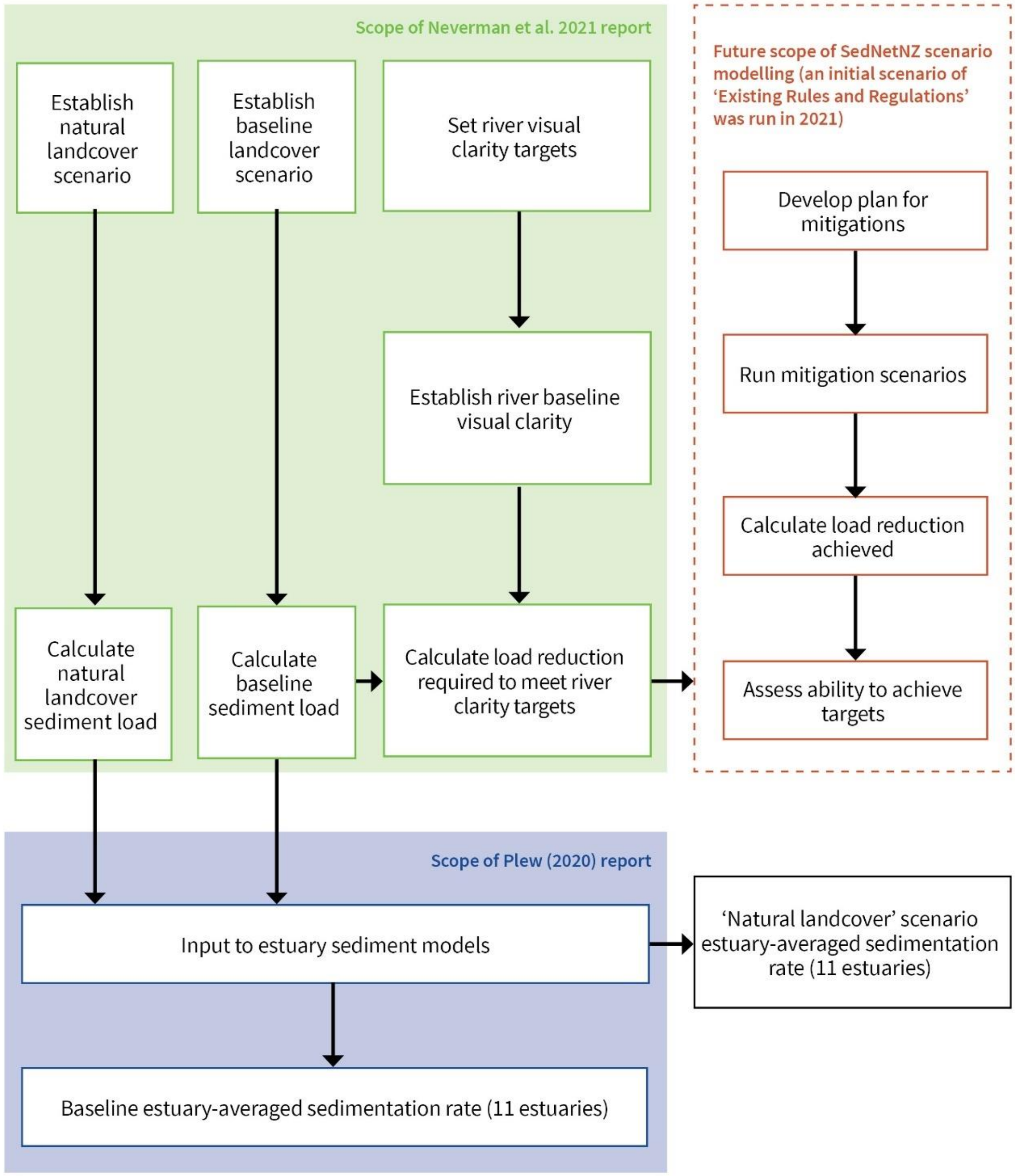


Figure 2: Sediment load reductions assessment components



(Modified from a presentation to stakeholders by Neverman et al., 22 June 2021)

Was the modelling process undertaken in accordance with general principles of best practice for modelling?

Inclusivity

24. Ideally environmental models would be conceived, designed, developed, evaluated, and applied in dialogue with people who live in and have a relationship with the area being modelled, and in a way that brings together modellers, technical experts, kaitiaki, and experts in maramataka, and mātauranga Māori.

25. Environment Southland has consulted widely with their community to develop draft freshwater objectives for the region. This has involved engaging with a regional forum and key stakeholders to inform the design of the council's approach to assessing the gap between the current state and what is required to the achieve draft freshwater objectives.

26. Throughout this process, Environment Southland has worked closely with Te Ao Marama Inc., who have been involved in both developing and deciding on the draft freshwater objectives for evaluation, and in commissioning the modelling work. A representative of Te Ao Marama Inc. explained to us they intend to use the outputs from this modelling work as an input to their own processes for assessing environmental state and for determining what is required to achieve the draft freshwater objectives.

27. While Te Ao Māori may not have been directly reflected in the technical modelling work *per se*, the perspectives and judgements of mana whenua influenced the setting of draft freshwater objectives, which established key benchmarks/targets/input parameters for the modelling. Having observed discussions between staff from Environment Southland and Te Ao Marama Inc., we consider there is:

    a. a high degree of trust between the organisations,

    b. comfort that Environment Southland's modelling team has a solid understanding of the interests and perspectives of mana whenua, and

    c. confidence that these perspectives have been accurately reflected in the way modelling work has been commissioned and project managed.

28. For these reasons, we consider the approach followed has been designed appropriately to allow Māori and non-Māori knowledge systems to grow together, work alongside each other – coordinating when it makes sense to do so and standing apart but alongside each other when that is appropriate.

Transparency and accessibility

29. In general, the methods used by the modelling teams were sufficiently clear, and data sources and assumptions were described in sufficient detail to facilitate interrogation and replication. In some instances, however, methodology was established through cross-referencing and reports were dense and technical. This required us to track back through referenced documents to be certain we understood the methods that were followed, the rationale for modellers' judgements, and the basis for choosing between different possible approaches.

30. It would have helped our review, as well as other audiences who sometimes have less specialised understanding, if the model reports, given they are linked, were prefaced by a summary document that clearly spelt out – in plain language – the:

    a. overall conceptual framework and modelling architecture (how the models link together, as we have attempted to show in **Figures 1 and 2**),

    b. methodology and assumptions followed by each modelling team,

    c. rationale for key choices made by the modellers, and

    d. key conclusions/findings of each modelling exercise.

31. It would also have helped if the summary document compiled links to the key pieces of work underlying the methods chosen and the key sources of data. This would make it easier for third parties to interrogate and review the foundations of the modelling projects.

32. We understand the reports were not necessarily intended to be public-facing and were intended to be read alongside summary/explanatory memos provided by Environment Southland. We also understand that budget and resource constraints due to the impact of COVID-19 restrictions limited the time available to report-writers and necessitated the cross-referencing referred to above.

33. It is also important to acknowledge that New Zealand's freshwater management framework has been in a statue of flux since the introduction of the first NPSFM in 2011. This is relevant because it generally takes several years to generate the information necessary to develop and train environmental models of the type under review. Environment Southland began this programme of work in 2018/2019 and, given the pace of evolution in New Zealand's freshwater management system, was forced to reorient its modelling programme 'mid-stream' in response to significant changes to policy settings, regulatory terminology, regulatory thresholds/criteria, and the scope of science required to comply with national direction. This constrained the amount of resource available for 'polishing' the reports and preparing them for communication to a wider public/lay audience. Given the dynamic context within which this modelling is taking place, it would have been helpful to create a table to keep track of changes in regulatory requirements, the implications of those changes for the modelling programme, and the nature of Environment Southland's responses to those changes.

34. We would like to note that the points we raise above are intended to be read as constructive suggestions to those undertaking future similar modelling processes in other regions and reflections for consideration by the project team at Murihiku/Southland.

Rigour

35. Much of our discussion focussed on the accuracy of model predictions, levels of uncertainty, and the implications of assumptions in the models' design. All the models under review have been developed in a data-poor context where there is a significant lack of understanding about core processes (e.g., contaminant attenuation and the influence of groundwater, sediment deposition in estuaries).

36. We note, for instance, that while groundwater contamination and any temporal lags between activities on land and their effects on water quality will have an influence on the freshwater objectives, this has not been factored explicitly into any of the modelling. Data

that are available in Murihiku/Southland suggest very short 'lag times' between activities taking place on land and their effect on water quality. We are mindful that data are sparse, and the assumption of an instantaneous link between land use and water quality is likely to be incorrect. In the future, should monitoring/observation determine that outcomes are different to what the models have predicted, it may be that the variance is due to incorrect assumptions regarding contaminant lag times rather than errors in the modelling itself. We note that Environment Southland is conducting further research into contaminant attenuation and the effect of groundwater on surface water quality. The outcomes of this research will have an important bearing on policymaking and limit setting under the NPSFM.

37. In some cases, the model reports describe what data are used but don't comment on the comprehensiveness or paucity of data and the implications of relying on assumptions or estimates derived from other locations as model inputs or to corroborate model outputs. In other cases, the implications of data-limitations are clearly spelt out. Overall, on close reading and having had the benefit of discussions with the report authors, it is evident that data and assumptions were analysed critically, and the judgements made by modelling teams were appropriate and defensible. In general, there appears to be enough data available to derive models suitable for undertaking a 'first-pass' assessment at the regional scale prior to stepping to more tailored modelling at finer levels of detail in areas of interest.

38. Ideally, given limitations in data, the model reports would contain a more comprehensive explanation of the source and nature of uncertainty (e.g., through further Monte Carlo analysis), tease out the implications of known unknowns, identify areas of potential unknown unknowns, and clearly spell out the implications of uncertainty for model users and for the potential scope of model application. Future projects should consider using diagrams to highlight areas of uncertainty and point out how they are likely to propagate through the model architecture.

39. Where information is incomplete and/or where there is uncertainty due to limits in the spatial or temporal scope of data it is sensible for councils to draw on other complementary sources of information to test, corroborate, and validate data collected in accordance with statutory and regulatory requirements (i.e., as mandated by the NPSFM and National Objectives Framework). By combining monitoring and modelling and by drawing on "multiple lines of evidence" to test and corroborate findings, council staff help ensure they are providing decision makers with the 'best available information' as required by the NPSFM. Future projects should consider more clearly describing any complementary work undertaken as part of their investigations.

40. We noted that, in some instances, there are sufficient data and there is background work that could be drawn on to facilitate more detailed modelling - including three-dimensional modelling in at least one, perhaps two, of the region's estuaries. Given the purpose of this modelling programme.[7] we consider it defensible that Environment Southland chose to prioritise uniformity of approach over opting for a mosaic of different models at different levels of complexity with potentially conflicting outcomes.

[7] See paragraphs 13 and 14 of this memo.

41. In reaching this conclusion we took into account advice from Environment Southland that, where comparative assessments were undertaken of the predictions of generalised versus detailed models, outputs from more detailed models tended to corroborate the findings of the generalised models, suggesting a reasonable level of concordance. Similarly, we took into account advice from Environment Southland that observational studies undertaken by the council tended to corroborate model predictions. In particular, we understand that observational studies tended to support the modelled prediction of widespread risk of periphyton growth in contrast to monitoring data which suggested widespread compliance with target attribute states (draft freshwater objectives) set for periphyton growth (i.e., there is a low risk of periphyton growth.)

42. Finally, we note that the project team within Environment Southland approached this task with obvious conscientiousness and integrity. We considered it a strength of the programme that team members ran their own internal evaluation process – pressing modellers on methodology and their rationale for favouring methods and model-types, and the basis for underpinning assumptions – and captured the outcome of this evaluation in concise and easy-to-understand memos tailored to a range of audiences.

Specific assessment

Snelder, T., (2021) *Assessment of Nutrient Load Reductions to Achieve Freshwater Objectives in the Rivers, Lakes and Estuaries of Southland Including Uncertainties: To inform the Southland Regional Forum process*, Land Water People

Is the modelling conceptually, technically, and scientifically robust?

43. We are generally confident that the modelling of nutrient concentrations, loads, and load reductions required to achieve freshwater objectives is robust. Overall, it identifies and acknowledges the key uncertainties, its scientific judgements are sound (e.g., adopting a middle ground scenario for stream shading), and it draws on the most robust dataset available for the purpose to which it is being put, and is scientifically sound.

44. Modelled estimates of the probability that periphyton biomass threshold objectives are complied with for different levels of spatial exceedance were the subject of considerable scrutiny. The panel noted that monitoring data shown on the Environment Southland website indicate monitored baseline states for periphyton mostly meet draft objectives (target attribute states) whereas the model's predictions suggested widespread exceedance of thresholds when 20 per cent and 30 per cent spatial exceedance criteria were applied.

45. After reviewing the model report and questioning the model author, the panel drew the following conclusions:

    a. The model and estimates of nutrient criteria are broadly consistent with Ministry for the Environment guidance for modelling nutrient-periphyton relationships.[8]

    b. The distribution of periphyton biomass in rivers can be inconsistent. We have imperfect understanding of the causes of variability in periphyton biomass and

[8] During our review the Ministry for the Environment released updated nutrient-periphyton criteria. We were aware of these changes, discussed their implications with the model author, and factored this change into our assessment of the modelling process. We consider it would help stakeholders' understanding if Environment Southland were to re-run the relevant models using these updated criteria, share the results with stakeholders, and make the model author available to explain the impact of this change.

imperfect indicators for predicting variability at a fine-grained level of spatial resolution – some degree of uncertainty in both measured estimates and model predictions are unavoidable.

c. The model's results are consistent with the results of similar models applied in the USA and EU – their R^2 values are similar, and the direction of response of biomass to each of the model's predictor variables is consistent with our general understanding of the controls on periphyton biomass.

d. The periphyton monitoring network in Southland is sparse (36 sites) and was designed to enable development of a regional periphyton model, which was subsequently superseded by, but continued to contribute to, national modelling. The monitoring was intended to capture conditions across a range of flow regimes and nutrient levels for the purpose of modelling biomass response, rather than to characterise regional conditions (i.e., the monitoring sites were not randomly chosen). The resulting dataset, while providing a reasonable estimate of periphyton biomass and variability at the suite of monitored sites, may not fully represent periphyton biomass patterns across the region.

e. Given the high levels of spatial and temporal variability in periphyton biomass, the model is not expected to perform well when predicting the incidence of periphyton at a specific location but is expected to perform better when predicting the likelihood at a large scale that a given proportion of sites drawn at random will exceed periphyton biomass criteria. As such, the model is designed to estimate the likelihood that a certain percentage of sites within a river network will experience excessive periphyton biomass rather than predict the presence or absence periphyton at a particular place and time.[9]

f. Observational studies conducted by Environment Southland show that excessive periphyton biomass does occur at certain times of the year so we are not surprised the modelling predicts there will be river reaches in Murihiku/Southland that exceed spatial exceedance criteria for periphyton. These periods of excess periphyton might be the most critical for impacts on values such as Ecosystem Health, Mahinga Kai or Recreation.

At what scale can the model(s) be used with confidence?

46. The model integrates the effect of nitrogen and phosphorus on different attributes across waterbodies and was designed to operate over a large geographic area. At high levels of aggregation, the model helps frame the scale and nature of water quality issues and gives a reasonably reliable indication of the magnitude reduction in contaminants required to achieve target attribute states ("draft freshwater objectives").

47. We are confident that model outputs are fit for the following purposes at the whole of region, 'freshwater management unit', and Environment Southland's 'reporting catchment' scales:

[9] See MfE Periphyton Guidance pp 10-13 for further explanation: Guidance-on-look-up-tables-for-setting-nutrient-targets-for-periphyton-010622.pdf (environment.govt.nz)

a. Estimating current (baseline attribute) state throughout the region, rather than just at monitoring sites.

b. Numeric estimation of load reductions required to reach target attribute states for pre-determined levels of spatial exceedance.

c. Numeric estimates of average instream nutrient concentrations.

48. In some instances, the model may be applicable at the sub-catchment level or areas upstream from critical points. This would only be the case where the area is sufficiently large and is well represented by observed data in the model training (i.e., several monitoring sites within that area).

49. At "finer-grained levels" – from sub-catchments to waterbodies and river segments – the model's predictions will struggle to account for local variation and are unlikely to provide a reliable indication of observed conditions.

50. In summary, data from monthly monitoring provide a reasonable estimate of conditions at the monitoring sites, but these patterns may not be representative of the whole region. In contrast, modelled estimates of baseline state provide a reasonable regional, or catchment, picture, but are less likely to correctly predict conditions at reach or site scale.

Is the modelling appropriate for the purpose Environment Southland intends to use it?

51. The modelling provides an evidence-based approach to the integrated management of catchments and estimating nutrient load reductions required across complex catchments to achieve target attribute states ("draft freshwater objectives") for multiple attributes in rivers, lakes, and estuaries. It represents a significant step forward in terms of implementing the NPSFM.

52. The modelling is closely linked to Environment Southland's monitoring data (i.e., it is data driven) which gives us more confidence in the predictive ability of the model for the purpose of estimating current contaminant loads at broad spatial scales. The modelling is not mechanistic, however, so has limited application in terms of assessing the likely effect of different potential mitigations or mitigation packages.

53. It is important to note there is significant uncertainty in the periphyton model that has been used to establish criteria for in-river nutrient concentrations. Periphyton growth is highly spatially variable, and the modelling is designed to identify the likelihood that local conditions will prompt biomass to exceed pre-determined spatial exceedance criteria. Model predictions regarding compliance with draft freshwater objectives are highly sensitive to the level at which exceedance criteria are set, which reflects how much risk the community is prepared to accept. Deciding what level of risk is "appropriate" is a policy rather than a science question and, given the modelling approach that has been taken, could conceivably lead to "under" or "over" protection of waterbodies. Exactly where exceedance criteria are set is a choice that sits with the council and local community – neither the modellers nor this panel have a view on where these criteria should be set, but we are confident that the model outputs provide a robust basis for informing policy decisions.

Plew, D., (2020) ***Models for evaluating impacts of nutrient and sediment loads to Southland Estuaries: To inform the Southland Regional Forum process*****, National Institute of Water & Atmospheric Research, NIWA CLIENT REPORT No: 2020216CH**

Is the modelling conceptually, technically, and scientifically robust?

54. The modelling described in this report was based on the estuarine trophic index and has a well understood theoretical underpinning. Local data were extremely limited, however, and where data were available there was a very low correlation between model predictions and observations of sedimentation in estuaries, with some significant outlying predictions. This could indicate that results from individual sampling points don't reflect average sedimentation across the estuary and suggests that further monitoring and model validation are required.

55. Overall, the modelling assumes a simple linear relationship between environmental pressures and environmental state. These relationships are very complex, site specific, spatially variable, difficult to define, and unlikely to be linear. That said, given the way in which the model outputs are to be used at this phase in Environment Southland's modelling process, and given there is not enough information to derive a more accurate relationship between pressure and state, we consider this assumption acceptable.

At what scale can the model(s) be used with confidence?

56. The modelling represented entire estuaries as a single basin which conceals variation in contaminant inputs from different contributing sub-catchments, or differences in sedimentation due to variable flushing rates, and variations in bathymetry and depth. This means the modelling will not accurately represent variability and/or capture the localised effects of sediment inputs to the estuaries.

57. Because the model averages sedimentation across an entire estuary, if the model is used to derive load reductions for contributing sub-catchments, there is a risk that some sub-catchments will have limits that are too stringent and some that are too lenient to achieve draft freshwater objectives.

58. Conversely, if the model's predictions are geared to be accurate at one monitored stream or river input rather than averaged across the estuary, load reductions derived from these estimates might be accurate for the sub-catchment upstream of the chosen input but wrong for the other contributing sub-catchments.

59. We understand that in Murihiku/Southland, while many estuaries have multiple river/stream inputs, contaminant loads in these estuaries tend to be dominated by inputs from single major rivers or streams. Ideally, the modelling should treat estuaries as multiple units if that is how they function, but we understand it may be difficult to determine exactly how estuaries behave and to represent this behaviour in models given the paucity of available data. In that regard, it is relevant that a more complex three-dimensional estuary model run in parallel to the uni-dimensional modelling in one estuary generated predictions that corroborated those of the simpler model.

60. We have noted earlier[10] the importance of using multiple lines of evidence to corroborate model predictions, particularly in a data-poor context. One additional line of evidence

[10] See paragraph 39 of this memo.

worthy of consideration in this context is historical aerial photography, which could be used to help infer longer-term estuary behaviour and complement/corroborate the results of sediment monitoring and modelling work.

Is the modelling appropriate for the purpose Environment Southland intends to use it?

61. The report was very transparent about the high levels of uncertainty associated with model predictions and how this limits the extent to which model outputs can be relied upon. The author's conclusions are appropriately tentative and there is a clear acknowledgement that more work is required before the models can be used in subsequent stages of Environment Southland's modelling programme.

62. Outputs from this model were used to derive a benchmark level of nitrogen that is likely to prompt the growth of phytoplankton or macroalgae, which was then used as an input by Snelder when estimating nutrient loads necessary to comply with target attribute states ("draft freshwater objectives). It was relevant to our consideration that subsequent modelling by Snelder indicated that in most instances the achievement of draft freshwater objectives is likely to be constrained by sensitivity to periphyton growth in rivers rather than the likelihood of phytoplankton or macroalgae growth in the estuaries at the 'end of the chain' of freshwater receiving environments.

63. We consider the model generates outputs that are appropriate for use at this early stage in Environment Southland's programme. While we agree with the model author that more work is required before the model can be used to establish limits, and that more sampling data are required to facilitate this, we would only consider this necessary in situations where an estuary is shown to be the critical limiting water body for limit-setting purposes.

Neverman A., Smith H., Herzig A., Basher L., (2021) *Modelling baseline suspended sediment loads and load reductions required to achieve Draft Freshwater Objectives for Southland*, Manaaki Whenua – Landcare Research, Contract Report: LC3749

Is the modelling conceptually, technically, and scientifically robust?

64. The modelling estimated suspended sediment loads necessary to achieve visual clarity objectives based on a relationship between increase in visual clarity and reduction in sediment load, which was published in reports that underpinned development of the sediment attributes in the NPS-FM.[11]

65. We note that (a) Environment Southland has very little long-term sediment load data, and (b) monthly fixed-interval water sampling typically misses most flow events that transport most of the suspended sediment, which are episodic processes driven by rainfall events. This means that Environment Southland's available sediment load data limits the ability to model sediment loads and their influence on visual clarity. The modellers responded to this lack of data by using SedNetNZ to predict mean annual sediment loads and modelled visual clarity data from Snelder for the baseline visual clarity and scenario. SedNetNZ is a mechanistic model which predicts sediment loads from input such as riverflows, land cover and slope. It was developed in Australia and is generally adapted on a catchment-by-catchment basis to the New Zealand context – in this case this included modelling to predict surficial and bank

[11] See Hicks et al 2019.

erosion. The bank erosion model is calibrated using data from Manawatu but we are unsure what approach was followed to calibrate or test the model's predictions of surficial erosion. We support the modellers' justification for adopting SedNetNZ – that the model is of intermediate complexity, that it represents erosion processes adequately, but that the number of parameters is small enough to be practical – and consider the decision to be defensible. But, without local calibration of the model's coefficients and in the absence of sensitivity analysis of these model components, we have limited confidence in the model's predictions of sediment load.

66. We noted that a mix of data from Fundamental Soils Layer (FSL) and SMap were used as inputs to the model. We understand the modellers took the view that it was preferable to represent the spatial variation in soil properties/erodibility by using the best available data rather than prioritise consistency and rely on one source of data. We are comfortable that the modellers' judgment in this regard is defensible.

At what scale can the models be used with confidence and is modelling appropriate for the purpose Environment Southland intends to use it?

67. The paucity of local data on sediment loads increases the uncertainty of model predictions, which magnified the uncertainty inherent in the regulatory framework for sediment management, given it is based on the relationship between increases in visual clarity and reductions in sediment load. These factors have forced the modellers to make some difficult choices.

68. We are satisfied the modellers approached their task with rigour, made use of the available data to provide decision-makers with the best available information, and have made defensible decisions. The levels of uncertainty associated with this modelling, however, increases the influence of normative (policy) judgements regarding where exceedance criteria thresholds are set – making it particularly important that decision makers are aware of this uncertainty and its implications when confirming objectives and setting limits.

69. We would like to emphasise two residual matters:

    a. We understand the rationale for using a long timeframe as this captures variation in sources over time such as the effect of episodic events linked to significant rainfall events (e.g., bank erosion and slips). But the SedNetNZ model appears not to account for land-use change, intensification, and implementation of mitigations over that 30-year period. Given our knowledge of the annual variation of water quality and trends over time, we remain to be convinced that a single averaged value over a 30-year period can reasonably represent a measured suspended sediment concentration metric when some key management factors will be changing over these 30 years.

    b. We understand there is often a strong relationship between the three parameters of Total Suspended Solids (TSS), Turbidity and Visual Clarity, but we also understand this relationship to be very location specific. In this case the modelling process in the Neverman et al 2021 report appears to have relied on a nationally derived TSS-Visual Clarity relationship from Hicks et al 2019, which may not accurately represent loads in the Murihiku/Southland context. This limits our confidence in the model's predictions at anything other than the 'whole of region' scale.

Ken Taylor **17 October 2022**

Science Review Panel Chair

Attachment 1: Prompts used to coordinate assessment

General principles	Prompts
Inclusive	• The range of perspectives and information sources incorporated into models • The role of Te Ao Māori and degree of focus on inter-generational health of Te Taiao.[12]
Transparent	• Degree of access to objectives, assumptions, sources of data and methods of data collection, mathematical frameworks employed, accuracy thresholds used, and quality-assurance processes followed • Extent to which limitations, and uncertainties, including any evidence gaps, complexities and areas of contention have been identified • The process used to ensure individuals and groups outside the project team (e.g. decision-makers and mana whenua, kaitiaki, policy, regulatory and operational staff in public authorities, parties likely affected by decisions made on the basis of model outputs) are able to feed in to evaluation processes, influence the design of the model, and comprehend its outputs and their implications.
Rigorous	• The quality and comprehensiveness of data • The appropriateness of quality assurance and evaluation processes (including planning, implementation, documentation, assessment, and reporting) to ensure the model and its components are suitable for its intended use and meet required/reasonable performance standards.
Accessible	• The extent to which model predictions and supporting analyses, model evaluation or peer-review reports, and model implications are easy to understand.
Specific elements	**Prompts**
Scientific basis	• The scientific theories that form the basis for models including their relationship to Te ao Māori, and extent to which they draw on mātauranga and maramataka.
Conceptual basis	• The attributes, relationships, and processes of the system relevant to the problem of interest.
Computational infrastructure	• The mathematical algorithms and approaches used in executing the model computations.
Assumptions and limitations	• The detailing of important assumptions used in developing or applying a computational model, as well as the resulting limitations that will affect the model's applicability.
Data availability and quality	• The availability and quality of monitoring and other data that can be used for both developing model input parameters and assessing model results.
Test cases	• The availability of basic model runs where an analytical solution is available, or an empirical solution is known with a high degree of confidence to ensure that algorithms and computational processes are implemented correctly.

[12] Ideally environmental models would be conceived, designed, developed, evaluated, and applied in dialogue with people who live in and have a relationship with the area being modelled, and in a way that brings together modellers, technical experts, kaitiaki, and experts in maramataka, and mātauranga Māori.

Corroboration (validation) of model results with observations	• Comparison of model results with data collected or observed in the field to assess the model's accuracy and improve its performance.
Benchmarking against other models	• Comparison of model results with other similar models.
Sensitivity and uncertainty analysis	• Investigation of the parameters or processes that drive model results, as well as the effects of lack of knowledge and other potential sources of error in the model. • Identification and explanation of implications for potential scope of model application.
Model resolution capabilities	• The level of disaggregation of processes and results in the model compared to the resolution needs from the problem statement or model application. The resolution includes the level of spatial, temporal, or other types of disaggregation.