Taihoro Nukurangi

Managing Upstream: Estuaries State and Values – Methods and Data Review

Stage 1B report

Prepared for The Ministry for the Environment

March 2018

Prepared by:
Anastasija Zaiko (Cawthron Institute)
Anna Berthelsen (Cawthron Institute)
Chris Cornelisen (Cawthron Institute)
Dana Clark (Cawthron Institute)
Richard Bulmer (NIWA)
Judi Hewitt (NIWA)
Leigh Stevens (Wriggle Coastal Management)
Rebecca Stott (NIWA)
Graham McBride (NIWA)
Chris Hickey (NIWA)
Jonathan Banks (Cawthron Institute)
Neale Hudson (NIWA)

For any information regarding this report please contact:

Neale Hudson
Environmental Chemist/Group Manager
Catchment Processes
+64-7-856 1772
neale.hudson@niwa.co.nz

National Institute of Water & Atmospheric Research Ltd
PO Box 11115
Hamilton 3251
Phone +64 7 856 7026

NIWA CLIENT REPORT No: 2017415HN
Report date: March 2018
NIWA Project: MFE16210

Quality Assurance Statement		
	Reviewed by:	Neale Hudson
A. Bartley	Formatting checked by:	Alison Bartley
	Approved for release by:	Dr David Roper

Contents

Tables

Figures

Executive summary

The "*Managing Upstream: Estuaries State and Values"* project aims to inform management decisions made when establishing freshwater objectives under the National Policy Statement-Freshwater Management. It is anticipated that the information provided will enable decisions to be made in a manner that better accounts for impacts on estuarine values. The project also aims to increase knowledge of the state of different estuary types in New Zealand. The technical work is being delivered for the Ministry for the Environment by an interdisciplinary team of researchers and scientists from Crown Research Institutes, several universities, several regional councils, and private consultancies.

This report builds on earlier work that identified and prioritised candidate attributes (variables used to inform upstream management) and state variables (indicators of the state of estuary values). It includes the identification and review of data likely to be useful for identifying critical attribute thresholds, and for providing baseline and reference information for state variables for three key estuary values: ecosystem health, human health and mahinga kai. Metadata on these data were compiled into tables and used to identify available data, information gaps, and to recommend the most important datasets for further use in the project.

Responses by estuary experts to an on-line survey were used as a form of validation (were the right variables were being targeted?), and to ensure that the limitations of measurement and analysis methods were documented. The survey captured areas of agreement (or disagreement) among experts regarding variables they considered to be most useful as attributes and/ or state variables to be captured, and commententary regarding monitoring of different variables (e.g., limitations, spatial or temporal considerations). In addition, the survey provided an opportunity for experts to comment on methods used to collect samples for analysis of variables, within a context of temporal and/ or spatial variability. Survey results and other information were compiled to create a series of methodology factsheets for the prioritised attributes. These factsheets describe gaps in datasets, along with methodological "bottlenecks". Bottlenecks included limitations or other factors that compromised the applicability of attributes over time, resolution, and/or space. This information helped determine the availability of data across regions and within specific estuary typologies.

A total of 36 groups of data (either compiled datasets or individual reports/data files) were identified for consideration in Stage 2 of this project, should it be commissioned. Of these, 19 were acquired for later project use; these represented 19 of the 27 prioritised attributes/ state variables. Except for state variables for *frequency of customary harvest closures* and *harvest area accessibility,* it is likely that additional variables will be covered in some capacity by the datasets identified. Most of the data identified, acquired or likely to be accessible, fulfilled Quality Assurance (QA) criteria. The Quality Assurance rating for remaining data remains uncertain, pending provision of further information.

From the gap analysis, we identified critical data shortages for several of the prioritised attributes and state variables. These included those relevant to human health and mahinga kai values such as shellfish metals, shellfish faecal indicator bacteria, and the distribution and abundance of shellfish. Some of these gaps may be addressed by datasets that were identified, but still need to be acquired or reviewed. In addition, some data gaps relate to variables that are likely to become increasingly important in the future (beyond the life of the project), such as emerging contaminants.

Information derived from the survey and assessment of various datasets was used to draw conclusions, make recommendations on the steps that will be required to address data gaps, and identify potential issues related to monitoring methods. Filling data gaps and addressing monitoring issues is scheduled for Stage 2 of this project. The proposed attribute variables most likely to benefit from further analysis, or from acquisition of additional data during Stage 2 cover all three of the ecosystem health, human health and mahinga kai estuary values. Proposed attribute variables for further development in Stage 2 include:

- Sediment deposition rate (measured and modelled).
- Water nutrients (total nitrogen (TN), total phosphorus (TP)).
- Total suspended solids.
- Water faecal indicator bacteria.
- Macroalgae.
- Macrofauna.
- Mud content/grain size.

The degree to which robust, standardised methods are used for both collection and analysis of these data is also important. The review identified methodological problems that are likely to be solved, either by implementing relatively trivial fixes to the data (e.g., conversion of values to common units) or following further data analysis. Both solutions will be implemented during Stage 2 of the project if commissioned. These solutions are related primarily to improved standardisation of methods and application of consistent quality assurance practices across datasets.

Temporal and spatial variability of the different variables was also considered to determine whether such variability is adequately understood, or whether it could limit the usefulness of a variable as either an attribute or a state variable. Experts agreed that more was known about large-scale temporal variability than large-scale spatial variability (i.e., differences driven by estuary type, between coasts and longitudinal gradients). This information will guide activities in Stage 2, acknowledging that some data-related issues may be 'fixed' or resolved through desktop analysis of existing datasets, while others may need to be addressed by collecting new data. Recommendations are made whereby data and knowledge gaps may be addressed in Stage 2, with emphasis on monitoring and data collection likely to be required to address gaps, and develop variables into attributes. Caveats regarding the sampling required to address issues associated with temporal and spatial variability of proposed attributes are provided in the detailed methodological factsheets. Addressing these caveats would assist with the development and implementation of attributes and state variables beyond the life of the project.

The project team acknowledges that many different approaches could have been trialled or followed to identify attributes and state variables and to then develop and recommend a prioritised list for further consideration. In view of the constraints imposed by time and resources, and because we had involved the overwhelming majority of New Zealand estuary experts in the process, we consider that the following steps taken were appropriate:

- A workshop was held to describe the process likely to be followed, present likely attributes and state variables, and elicit feedback.

- Expert opinion and feedback was actively sought using easy-to-complete on-line surveys; survey responses provided expert opinion regarding likely attribute and state variable selection, and prioritisation of attributes and state variables.
- A larger number of attributes and state variables than ideal were included through the selection and prioritisation process (erring on the side of caution) to ensure that likely candidates were not eliminated prematurely.
- Expert opinion via surveys was also used to direct the gap assessment, assist with quality assurance and desciption of monitoring strategies and protocols.

The list of selected and prioritised attributes and state variables reflects the overwhelming preference of as many responses as were received. What is presented was arrived at through a transparent process, is considered technically defensible, and will enable the Ministry to make robust decisions during subsequent Stages of the project.

1 Introduction

The Ministry for the Environment (MfE) and regional councils have recognised that when setting management objectives and freshwater limits under the National Policy Statement-Freshwater Management (NPS-FM), there is also a requirement to protect estuary values, which is a logical requirement for integrated catchment management. The *"Managing Upstream: Estuaries State and Values"* project aims to provide the science to understand the impacts that limit-setting in freshwater management may have on estuarine values. This information will in turn enable future management decisions made regarding freshwater inputs into estuaries to be consistent with or support estuary values.

The technical work to provide this underpinning science is being delivered for MfE by an interdisciplinary team of researchers and scientists from NIWA, Cawthron, Universities (Auckland, Canterbury, Otago and Waikato), independent consultancies (Wriggle Coastal Management, Streamlined Environmental Ltd), Landcare Research Limited and several representative regional councils (Auckland, Bay of Plenty, Hawke's Bay, Waikato, and Southland)[1]. Regional council representatives are included in the project team as part of an Estuaries Partners Group (EPG) that contributes to relevant aspects of the project and provides feedback on report outputs.

The project aims to provide the scientific information required to:

- help inform management decisions made when establishing freshwater objectives under the National Policy Statement for Freshwater Management 2014 (NPS-FM), and
- increase knowledge on the state of different estuary types in NZ.

The project comprises three stages:

- Stage 1 (currently underway), includes the following activities:
  - identification of attributes and important indicators of estuarine state (the focus of the Stage 1A Report; Cornelisen et al. 2017)
  - review of available data and monitoring methods (included in this report)
  - identification of gaps in data and monitoring methods that limit full development of estuarine attributes required to manage freshwater limit-setting (included in this report), and
  - provision of advice regarding further development of attributes, state variables and monitoring protocols (included in this report).
- Stage 2 would include the following activities:
  - identification of critical thresholds for estuarine attributes that will be required to establish freshwater limits
  - provision of baseline and reference information to aid in the monitoring and assessment of estuarine state, and

[1] Two regional council representatives have specific mandate to inform other regional councils collectively, and to provide feedback on project delivery from the regional council perspective.

    - establishment of standardised monitoring protocols likely to enable adaptive management approaches for addressing upstream pressures on estuaries.

- Stage 3 (if commissioned) is likely to include the development of tools to assist with making management decisions, such as frameworks for limit setting.

All three stages of the project focus on three national-level values identified by MfE, that are common to all estuaries:

- ecosystem health
- human health for recreation, and
- mahinga kai.

1.1 Terminology and abbreviations

Key terms used in this project include **value**, **attribute**, and **state variable** (Table 1-1). Additional information on these and other terms were provided in the Stage 1A report, and a comprehensive glossary of terms with brief definitions related to the project is provided in Section 6 of this report.

Table 1-1: Definitions of terms.

Term	Definition	Example(s)
Value	Intrinsic qualities, uses or potential uses associated with estuaries. They may be qualities or uses that people and communities appreciate about estuaries and wish to see recognised (maintained or enhanced).	Shellfish gathering, bird watching, and swimming. Intrinsic values include ecosystem health, which encompasses the maintenance of ecosystem functions, natural form and character, and the provision of ecosystem goods and services.
Attribute	Measurable variables, including physical, chemical and/or biological properties that are directly affected by upstream aspects to be managed, such as sediments and nutrients. Attributes must be manageable, and directly support values.	A measure of mud content in the estuary, which is closely linked to sediment loading in the catchment.
State variable	Measurable variables (or composite metric of multiple variables) that provide information about the condition, or state, of an estuary value. State variables are useful for reporting and communicating the change in estuary condition over time in relation to the value.	The areal extent of seagrass, the diversity of macrofauna, or the frequency of shellfish harvest closures in an estuary.
Aspect to be managed	Aspects of catchments that need to be managed in order to maintain and enhance estuary values.	Loading of nutrients, sediments, faecal bacteria, as well as other contaminants and toxicants (such as metals and emerging contaminants).

The project focuses on three values of national relevance identified by MfE that apply across all estuaries, namely: **ecosystem health, human health for recreation**, and **mahinga kai** (Table 1-2). Ecosystem health and human health for recreation are also 'compulsory national values' for fresh water, and these are considered compulsory for councils to include in objective setting when implementing the NPS-FM.

Table 1-2: National values for estuaries. Text modified from that used for the NPS-FM.

Value	Definition	Aspects to manage
Ecosystem health	The ability of an estuary to support an ecosystem appropriate to its type. In a healthy estuary ecosystem, ecological processes are maintained, a range and diversity of indigenous flora and fauna occur, and there is resilience to negative change.	Loading of nutrients, sediments, toxicants such as heavy metals from stormwater runoff, and habitat loss.
Human health for recreation	Recreation in estuaries ranges from activities involving full immersion, such as swimming and diving, to those with less contact with the water, such as boating. The suitability of an estuary for water-based recreation depends, among other things, on whether water quality will adversely affect human health.	Loading of faecal contaminants including pathogens (viruses and parasites), as well as loading of toxicants such as heavy metals and emerging contaminants (e.g., those associated with pharmaceuticals, petroleum products).
Mahinga kai	Māori traditional food species gathered from the environment. The definition includes the places where these species are gathered and the practices involved in their collection. Indigenous estuarine species have traditionally been used as food, tools, or other resources. The inter-generational transfer of knowledge and practices related to mahinga kai is an important means of maintaining iwi traditions.	Aspects to be managed for mahinga kai overlap with those for ecosystem health and human health for recreation. Mahinga kai requires sustainable populations of kai species, which depend on a healthy ecosystem, and the ability to harvest and consume kai requires the loading of contaminants that affect human health to be managed.

Attributes ultimately provide the link for transforming values and high level narrative objectives into numeric objectives which in turn provide for defining limits and management actions. Essential criteria for attributes include their ability to:

- link to the values
- be manageable through freshwater inputs
- be measurable and predictable, and
- set management objectives.

State variables are used to assess (and report on) estuary condition and the state of estuary values, and must respond (at least in part) to changes in upstream pressures. In addition to variables that serve as attributes and state variables, supplementary variables need to be monitored to assist in interpreting the data provided for other variables. For example, water temperature and salinity may not serve as attributes or state variables, but are useful when interpreting results and understanding drivers of change.

Following on from the Stage 1A report, we use the broad estuary classification scheme developed as part of the Estuary Trophic Index toolbox (Table 1-3), following recommendations provided by Cornelisen et al. (2017). Throughout the text we refer to several councils using the abbreviations in Table 1-4.

Table 1-3: Estuary typology classification developed by Robertson et al. (2016) and used in this report. The types were assigned according to the Coastal Explorer tool.

Abbreviation	Abbreviation used in this report
ICOLL[2]	Intermittently closed/open lakes and lagoons estuaries.
SIDE	Shallow intertidal dominated estuaries.
SSRTRE	Shallow, short residence time tidal river and tidal river with adjoining lagoon estuaries.
DSDE	Deeper subtidal dominated, longer residence time estuaries.
Unknown	Unknown or uncertain type (no entry in the Coastal Explorer and no typology information available from the metadata).

Table 1-4: Regional council names and abbreviations used in the report.

Abbreviation	Regional Councils
AC	Auckland Council
BOPRC	Bay of Plenty Regional Council
ECAN	Environment Canterbury
ES	Environment Southland
GDC	Gisborne District Council
GWRC	Greater Wellington Regional Council
HBRC	Hawke's Bay Regional Council
HRC	Horizons Regional Council
MDC	Marlborough District Council
NCC	Nelson City Council
NRC	Northland Regional Council
ORC	Otago Regional Council
TDC	Tasman District Council
TRC	Taranaki Regional Council
WCRC	West Coast Regional Council
WRC	Waikato Regional Council

[2] ICOLLs are now considered a subcategory of both SIDE and SSRTREs (to better reflect their modifying nature on those estuary types).

1.2 Report aims and scope

Building on the Stage 1A work, this second report aims to identify and review existing methods for monitoring attributes and state variables, while also focussing on the identification, acquisition and review of datasets most likely to be useful in Stage 2 of the project. This includes identifying gaps in data, and limitations in monitoring methods that may need to be addressed in order to develop effective attributes and state variables.

The scope for this report was set by the list of candidate attributes and state variables from the Stage 1A report. The Stage 1A report identified a set of variables that met the criteria for an attribute, and have the greatest potential to be used to manage upstream pressures affecting the three national-level estuary values. The three values, and aspects to be managed that impact on these values, are each represented by a number of underlying candidate attributes, as indicated in Figure 1-1.

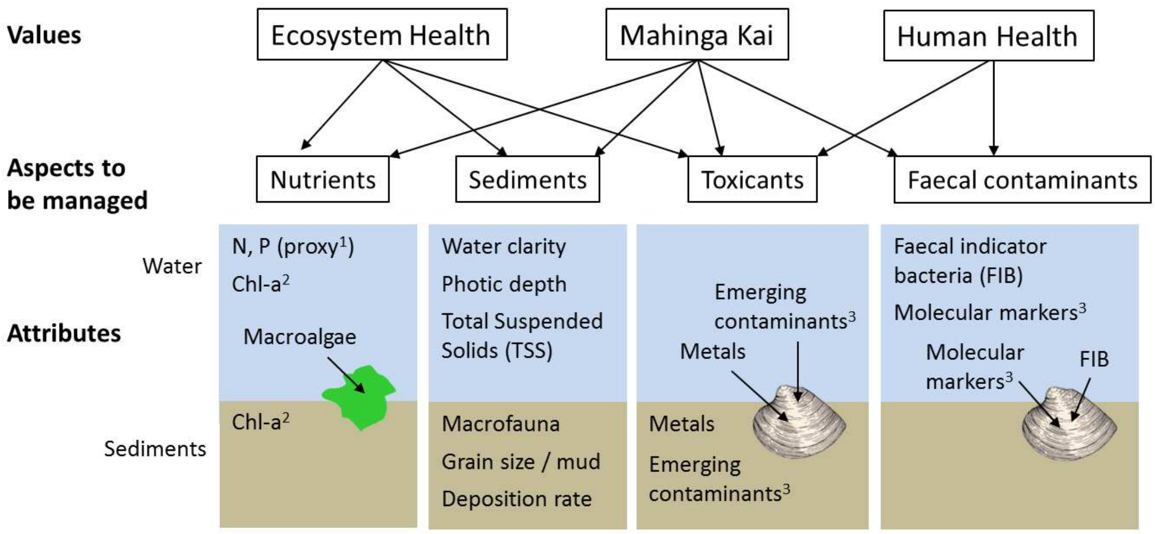


Figure 1-1. Variables recommended for further consideration as attributes in the Stage 1A report. [1]For nutrients such as nitrogen (N) and phosphorus (P), a proxy, such as modelled potential nutrient concentrations may be used. [2]Chl-a is a proxy for phytoplankton in the water and microphytobenthos (small algae) in the sediments. [3]The inclusion of emerging contaminants and molecular markers for faecal bacteria and pathogens is intended to mark their potential role in managing and monitoring estuaries following further research and development. It is unlikely these would be developed into attributes within this project.

This report focuses on assessing methods and datasets related to the list of variables in Figure 1-1. Realistically, only three to six variables that capture both the values and aspects to be managed are likely to be fully developed into attributes within the three stages of this project, with perhaps a slightly larger number of state variables. Taking this into account, and considering the essential criteria and role of attributes (see Section 1), as well as further information on methods and data, we aim to further refine the candidate attribute list to focus Stage 2 efforts.

The Stage 1A report also developed a list of candidate state variables likely to provide information about the condition, or 'state', of New Zealand's estuaries and values. In Table 1-5, these variables are arranged according to the three values and key categories that need to be addressed in order to support the value. This report also considered methods and data for these candidate state variables. While not all will serve as state variables, the majority of the variables listed already contribute to estuary monitoring programmes and could serve as 'supplementary' variables, used to aid interpretation of data collected for attributes and state variables.

Table 1-5. Variables recommended for further consideration as state variables. Those **bolded** are also candidate attributes.

Value	Category	Recommended priority variables
Ecosystem Health	Water quality	**Nutrient concentrations (N, P)** (can be modelled estimates) **Chl-a** Dissolved oxygen **Water clarity (e.g., Secchi disk)** **Total suspended sediments (or consideration of proxy such as turbidity)**
	Sediment quality	Broadscale extent of dominant substrate types, including: ▪ areal extent of mud ▪ areal extent of anoxic bottoms **Rate of sediment deposition** Fine-scale sediment variables at select sites, including: ▪ **grain size / mud content** ▪ sediment nutrients ▪ Total Organic Carbon (TOC) ▪ sulphides ▪ redox potential discontinuity (RPD) ▪ **sediment metals** ▪ **chl-a**
	Habitat quality and diversity	**Macroalgae: OMBT EQR from ETI toolbox** Broadscale extent of habitats, including for example: ▪ areal extent of seagrass ▪ areal extent of opportunistic macrolgae ▪ areal extent of salt marsh ▪ areal extent of shellfish beds ▪ areal extent of dominant substrate types
	Species diversity	**Macrofauna variables** (includes shellfish)
Human Health for Recreation	Bathing water quality	**Faecal indicator bacteria (FIB)** Frequency of bathing beach closures
	Shellfish quality	**Faecal indicator bacteria (FIB) in shellfish** Frequency of harvest closures (recreational & commercial) **Metals in shellfish**
Mahinga Kai	Shellfish	Shellfish distribution and abundance
		Frequency of customary harvest closures
		Harvest area accessibility
	Finfish	Finfish diversity and abundance

In this report, we focus on available data and methods available for measuring variables in estuaries while excluding those for upstream freshwater variables. For further development of attributes, measured attribute data must have meaningful relationships with upstream aspects to be managed (such as sediment accumulation within the estuary and the inflow sediment load). Establishing these relationships and identifying thresholds (e.g., upstream limits linked to bands for attributes) will require relevant data derived from catchment inflows.

We have assumed that datasets representing upstream aspects to be managed, such as suspended sediment or nutrient concentrations and loads, have been reviewed through work relating to the NPS-FM, and in many cases have been compiled (e.g., in the Land Air Water Aotearoa (LAWA) on-line data system, Table 1-6). In addition, we anticipate that these data, along with existing tools for estimating upstream loads (e.g., CLUES[3]), will increasingly be used when developing the bands or thresholds for attributes (Stage 2), and for the development of tools for assisting with the implementation of estuary attributes to inform freshwater management (Stage 3).

Table 1-6: Overview of existing/required data for establishing relationships between prioritised attributes and relevant upstream aspects to be managed. *Potential attributes for future development.

Prioritised attributes	Aspects to be managed upstream	Existing/required data from upstream
Water nutrients (TN, TP)	Nutrient enrichment	Nutrient concentrations and loads measured in rivers, or load estimates derived from tools such as CLUES.
Water chlorophyll-a (Chl-a)	Nutrient enrichment	
Water visual clarity	Sediment loading, nutrient enrichment	River total suspended sediment (TSS) concentrations, estimates of catchment sediment loading from existing tools.
Total suspended sediment	Sediment loading	
Water faecal indicator bacteria (FIB)	Faecal contaminants	River FIB concentrations, and possibly point source discharge concentrations. Loading from existing tools such as CLUES.
Macroalgae	Nutrient enrichment	Nutrient concentrations in rivers or loading estimates from tools such as CLUES.
Sediment Chl-a	Nutrient enrichment	
Macrofauna	Sediment loading, contaminants, nutrient enrichment	As above for sediment and nutrients.
Mud content/grain size	Sediment loading	Data on TSS and grain size in rivers. Integrated catchment runoff and hydrodynamic-sediment transport models.
Sediment deposition rate	Sediment loading	
Sediment metals	Toxicants	Metal concentrations in rivers, outfalls and stormwater.
Shellfish metals	Toxicants	
Shellfish faecal indicator bacteria (FIB)	Faecal contamination	As for Water FIB.
*Emerging contaminants (ECs)**	Toxicants,	EC concentrations within rivers, outfalls and stormwater.
*Molecular FIB markers**	Faecal contamination	As for Water FIB, and molecular markers in rivers.

[3] https://www.niwa.co.nz/freshwater-and-estuaries/our-services/catchment-modelling/clues-catchment-land-use-for-environmental-sustainability-model and https://www.niwa.co.nz/freshwater-and-estuaries/research-projects/estuarine-water-quality-the-clues-estuary-tool

2 Approach to data identification and review of attributes and state variables

For candidate attributes and state variables identified in the Stage 1A report, we describe the steps that were followed to review:

- data most suited for baseline assessment
- short- and long-term strategies for monitoring identified attributes and state variables, as well as
- gaps in data and methodological bottlenecks that could prevent use of the variables as either attributes or state variables.

Key steps (outlined in Figure 2-1), included:

- Identifying and collating information on existing monitoring methods that may be used to obtain data on the prioritised attributes and state variables.
- Identifying and collating available datasets (including metadata information), for each of the attributes and state variables.
- Undertaking a degree of quality assurance (QA) on the datasets.
- Using an online survey to canvas expert opinion regarding:
  - the prioritisation of attributes and state variables (validation/sense-checking)
  - gauging certainty of knowledge around variability, predictability of attributes and state variables, and
  - identification of potential issues that may constrain further development.
- Assessing methods and existing data for consistency and limitations, including gaps in knowledge and data.
- Providing recommendations to assist in planning work required in Stage 2 of this project (if commissioned); these included:
  - recommendations around monitoring methods that may require further development, for both attributes and state variables
  - identifying major gaps in data or knowledge to be addressed.

The final recommendations were revised in response to comments provided by MfE, external reviewers, including agreed outcomes of discussions with experts and project partners at the project closure meeting. This led to further shortening of the list of candidate attributes in order to focus efforts in Stage 2.

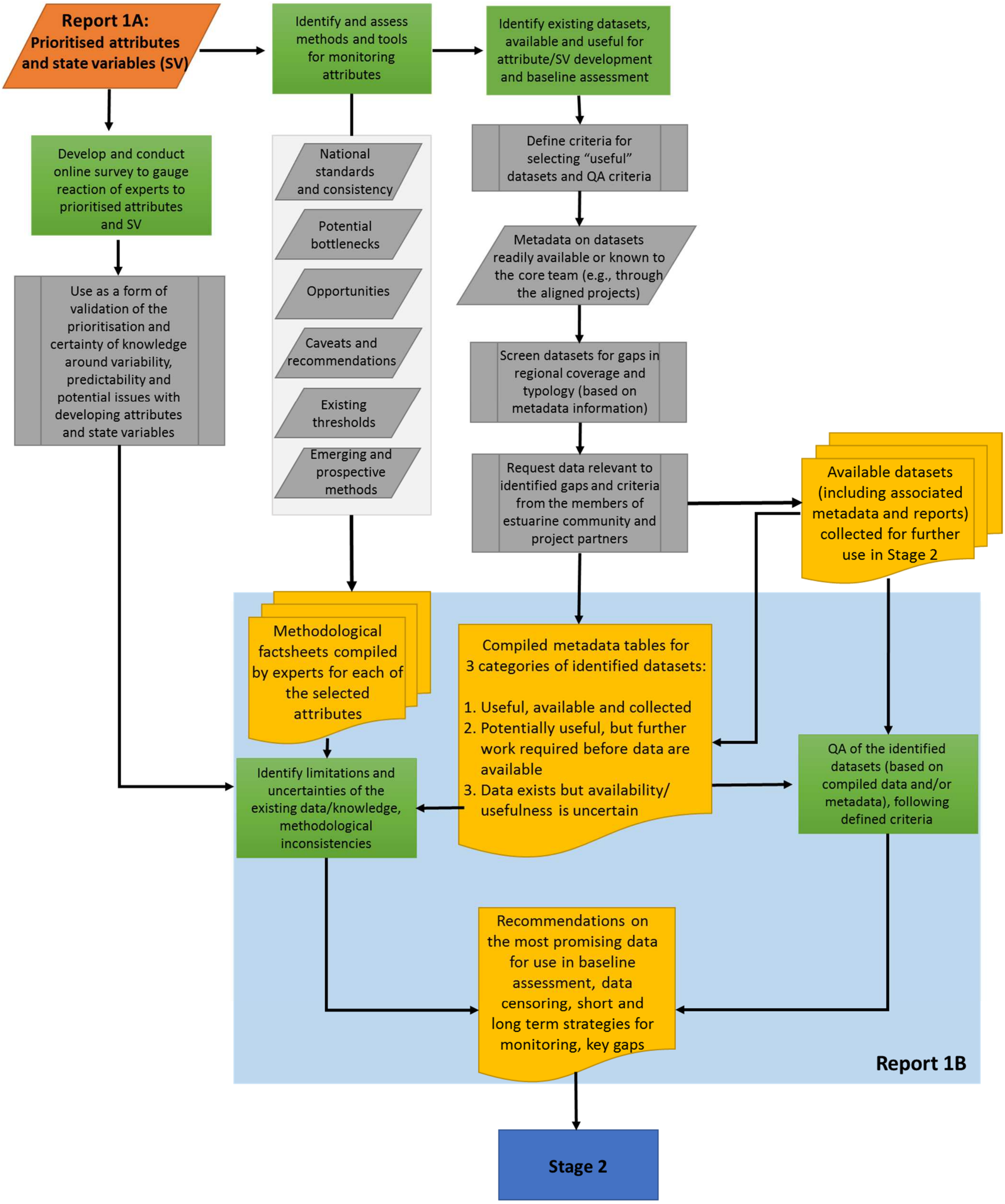


Figure 2-1: Schematic of workflow for this phase of the project.

2.1 Methods for monitoring and analysing attributes and state variables

We utilised expert knowledge to critically review monitoring and analytical methods with the objective of gauging differences of opinion, and identifying real or perceived problems associated with the list of attributes and state variables that had been prioritised for future use. The outcomes of this review were summarised in detailed factsheets that described the methods used, the consistency of methods of data collection and analysis in current use, and potential bottlenecks preventing easy use of candidate variables as either attributes or state variables (Appendix A).

In the introductory section of each factsheet, a brief rationale is provided to explain why the variable is a promising attribute for upstream management of New Zealand estuaries, and to indicate its suitability for assessing estuarine health with regard to three national-level values (ecosystem health, human health for recreation, and mahinga kai). The methodological overview in each factsheet summarises information derived primarily from current New Zealand monitoring- and research programmes, as well as from other national initiatives, while also considering international practice. An overview considered four major steps associated with acquiring environmental data for a variable:

- **Sampling design** – site selection, spatial extent of sampling, sample replication, etc.
- **Sampling procedures** – methods used to collect samples in the field, sample preservation, etc.
- **Laboratory analyses** – procedures used for deriving raw data (where applicable).
- **Computational approaches and derived metrics** – methods used for deriving the final attribute values, and metrics used to report on the state of the variable (where applicable).

For each of the steps above, we considered the following information:

- **National standards/ guidelines and consistency** – whether national standards (e.g., National Environmental Monitoring Standards, NEMS) and/or standardised protocols/guidelines for monitoring an attribute exist (providing references to relevant publications), and an overall evaluation of method consistency at national level.
- **Potential bottlenecks** – major limitations likely to impede attribute development, limit or compromise data quality, resolution and applicability of an attribute over time and space.
- **Opportunities** – opportunities for optimising or improving attribute monitoring.
- **Caveats and recommendations** – requirements for:
  - obtaining meaningful and robust information on the attribute in a nationally consistent and cost-effective manner (also considering temporal frequency and spatial extent), and
  - deriving robust thresholds for management of upstream water quality to meet the purposes of Stage 2 of this project.

The factsheets provide information on threshold values of an attribute (where available), and make recommendations for establishment of threshold values in the context of managing freshwater inflows to support estuary values. An overview of prospective methods proposed for improved monitoring of attributes is also provided in these factsheets.

Methods for monitoring state variables that were not also considered as candidate attributes were compiled in a summary table (Appendix B). In most cases, methods for monitoring these are included in existing estuary monitoring protocols (e.g., the Estuary Monitoring Protocol (EMP; Robertson et al. 2002), and the Estuary Trophic Index toolbox (ETI; Robertson et al. 2016a,b).

2.2 Identification and acquisition of datasets

The following sections describe the actions taken to identify 'available' and 'useful' data, including those required to fill gaps associated with regional coverage and according to estuary typology. Collation of these data and assessment of methodologies for consistency and quality is also described.

2.2.1 Criteria for selecting potentially useful datasets

For the candidate attributes and state variables (see Section 1.2), we targeted data that were **available** and **likely to be useful** in Stage 2 of the project for:

- identifying critical thresholds for attributes, and
- providing baseline and reference information for state variables.

Criteria used for determining data 'usefulness' were applied to fine-scale[4] data, which are dependent on a discrete sampling approach. These criteria included:

- **Sampling coverage and replication:** data were collected at more than one location in an estuary (>~30 m apart), or at more than one time (>~1 month apart, or calm vs storm conditions), or replicated spatially (5 or more samples) or temporally (e.g., tidal state).
- **Site description**: information is available to indicate site representativeness (i.e., habitat type, tidal position, susceptibility to upstream pressures, etc.).
- **Method documentation:** full description of sampling and laboratory methods were available, including quality assurance (QA) processes.
- **Data availability:** data can be made available to the project.

For broad-scale data[5] (e.g., extent of dominant substrate types, extent of habitats), potentially useful datasets were identified in consultation with core team members, and after considering available information regarding:

- **Required spatial data:** e.g., areal extent of mud, seagrass, mangroves, macroalgae, intertidal area, subtidal area, etc.
- **Sampling event:** date of sampling and tidal state at time of sampling for areal variables.
- **Spatial accuracy and field validation.**

[4] "Fine-scale" refers to site specific variables such as sediment grain size or macrofauna variables.
[5] "Broad-scale" refers to spatial variables such as areal extent of dominant substrates or habitat types.

- **Grain size verification of substrate.**
- **Data availability.**

2.2.2 Metadata collection and preliminary gap analysis

We first undertook a gap analysis for the fine-scale (versus broad-scale) attributes and state variables to identify data we would target for acquisition. This analysis used previously identified, readily available datasets likely to have future use within the project. The Oranga Taiao Oranga Tāngata (OTOT) dataset[6] and the MfE water quality dataset (Dudley et al. 2017) were used for the preliminary gaps analysis as they provided the most comprehensive regional coverage for benthic (OTOT) and water column (MfE) attributes/state variables. We screened these datasets to identify gaps in regional and typology coverage. Metadata recently collected from regional councils for identifying coastal data were also considered (Bolton-Ritchie and Lawton 2017); in future some of these data may be included in the Land Air Water Aotearoa (LAWA[7]) on-line data system.

We then contacted the core team and research partners, as well as staff at other regional and unitary councils by email, requesting provision of metadata for data that met the 'usefulness' criteria identified in section 2.2.1, and which could probably be used to fill identified gaps. Requested metadata included:

- name of the dataset/project
- timeframe of the data
- parameters covered
- regions covered
- number of estuaries (hydrosystems) covered
- whether permission or a data sharing agreement was likely to be required prior to use of these data.

We undertook a metadata collection and gaps analysis for broad-scale attributes/state variables (e.g., 'broad-scale extent of dominant substrates' and 'broad-scale extent of habitats') using a similar process.

2.2.3 Targeted collection of data to address identified gaps

Using the criteria identified above, we screened the acquired metadata to identify data likely to be useful for this project. These data, and additional metadata, were then requested from the estuarine scientific community and project partners. Data were provided in several formats, including data files (e.g., MS Excel and csv files), reports (as PDF and MS Word documents), website links and emails.

Not all data requested were received - in these cases, it was usually because the contact person was unable to respond (i.e., away on leave/ overseas). If considered useful, these data will be requested again as part of Stage 2 of the project. Some of the received data overlapped with datasets

[6] https://www.mtm.ac.nz/oranga-taiao-oranga-tangata/ and Berthelsen A, Goodwin E, Atalah J, Clark J (in prep) User manual for a national estuary dataset. Cawthron Institute.
[7] https://www.lawa.org.nz/

previously identified as useful (e.g., OTOT and Wriggle datasets). These were reviewed but not included in the final set of data recommended for further use.

2.2.4 Compilation of metadata tables for identified datasets

Metadata for 'useful' datasets were compiled into a table which identified the attributes/state variables within each dataset. The datasets were divided into three categories according to their current availability (also shown in Figure 2-1):

Category 1: Data are useful, available and was provided for use in this project along with accompanying metadata.

Category 2: Data are potentially useful and existing, can be made available later.

Category 3: Data exist but future availablility and/ or usefulness is uncertain.

All relevant files were downloaded and saved for future use. For some variables, information regarding the state and availability of datasets were not readily available. These included compilations of the seagrass spatial extent data (currently in preparation by Department of Conservation). [8]

2.3 Quality assurance of acquired datasets

A further step in the evaluation of the potential 'usefulness' of each dataset in Stage 2 of the project involved application of a quality assurance (QA) process. We first identified and defined QA criteria according to the methods used for collection and analysis of the data, as well as data entry (Table 2-1). For fine-scale data this included information on the site, including habitat type and representativeness; such information would be necessary if data from the site were to be used in determining thresholds for attributes or analysing estuarine state. We then evaluated each attribute/state variable within each dataset according to these criteria, and compiled the results of this process into a QA table. Where QA criteria information was unknown (usually due to insufficient metadata), the term 'to be confirmed' (TBC) was used in the relevant section of the QA table. Once the initial QA evaluation was completed, we made a concerted effort to gather the missing information from key contacts for all Category 1 datasets - these efforts were generally successful.

The Yes/No answers, and additional comments provided for the QA criteria for each attribute / state variable above were then assessed to determine an overall QA rating (Table 2-1). The QA rating provided an indication of whether the data were suitable for use in Stage 2 of the project.

[8] Helen Kettles pers. com.

Table 2-1: Quality assurance (QA) criteria used to evaluate the data. The overall QA rating is shown in bold at the bottom of the table. Compliance with the QA criteria was assessed at dichotomous scale (Yes or No).

QA criterion	Explanation
Site names	Sampling site names reported.
Units	Measurement units reported.
Reasonable temporal-spatial replication (applies to fine-scale data only)	More than one location in the estuary (~>30m apart), or at more than one time (~>1mo or calm vs storm), or be spatially (5 or more samples) or temporally (tidal state) replicated.
Timeframes	Collected within the last 10 years (applies to state variables only).
Site descriptions (applies to fine-scale data only)	Site descriptions available, including representativeness of sites (i.e., habitat type, tidal position etc.).
Methods descriptions	Full description of sampling and laboratory methods available
Sample collection method	A standard method was used. (If 'Yes', additionally noted whether the method consistent with EMP or Dudley et al. or ETI - for the broadscale data).
Sample analysis	A standard method was used. (If 'Yes', additionally noted whether the method consistent with EMP or Dudley et al. or ETI - for the broadscale data).
QA	Was the analysis QA'd? Was the data entry QA'd?
Data Entry (areal variables)	Is there information about the date of sampling and tidal state?
Data Entry (sediment point sampling)	Is there information about the date of sampling, number of replicates, site position (including representativeness information and location relative to tide) and site extent?
Data Entry (water column point sampling)	Is there information about the date of sampling, tidal state, number of replicates, site position (including representativeness information), water depth and site extent?
Data Entry (general)	Check completed for random errors (check highest and lowest values)
Recommendations for data censoring	**Are all QA criteria met?**

2.4 Expert survey

An online survey was used to validate the prioritisation of attributes/ state variables, to assess certainty of knowledge regarding variability and predictability, and to determine whether issues related to developing attributes and state variables were anticipated. It comprised two parts:

- Part A – to gauge the reaction of experts (project partners) regarding the selection of attributes and state variables[9].
- Part B – to address variability and predictability, and potential issues associated with prioritised variables that need to be addressed in the development of an attribute or state variable. Project partners and other technical experts answered questions about variables associated with their area(s) of technical expertise.

Respondents were encouraged to answer only those questions within their area of expertise. In Part B, a dichotomous scale ('Yes'/'No') was provided for respondents, avoiding neutral or unsure responses. *Post-hoc* review of responses was performed to confirm the eligibility of answers (by ensuring that expertise-related responses were received).

Free text comments fields were provided to capture general comments, concerns and variable-specific recommendations regarding:

- a particular region and/ or depth in the estuary where measurement of the variable would be required to make it a good attribute and/ or state variable
- a particular tidal state under which the variable should be measured to make it a good attribute and/ or state variable
- a particular time of year during which the variable should be measured to make it a good attribute and/ or state variable
- potential problems with the variable likely to make it unsuitable (at present) as an attribute and/ or state variable.

2.5 Assessment of data gaps and identifying the data most promising for further use

Analysis of acquired datasets and supporting metadata was undertaken to determine gaps, as well as the availability of data across regions and within specific estuary typologies. Information collected in the methodology factsheets and survey responses was used for identifying existing methodological bottlenecks in available datasets (i.e., limitation or compromise of applicability over time, resolution, and/or space), and to prioritise the potential attributes/ state variables.

Based on regional, typology, variable coverage and QA results, data were ranked in terms of potential importance by applying the sum of the following scores:

- if number of estuaries covered >1 = 1
- if temporal data (> 5 years) =1

[9] Molecular markers for faecal contamination and emerging contaminants were not considered in the survey, given they were suggested as potential variables for future consideration in Report 1).

- types of estuaries covered:
  - 2 types = 1
  - 3 types = 2
- number of regions covered:
  - 2 regions = 1
  - > 2 regions = 2
- number of variables covered:
  - 2 variables = 1
  - >2 variables = 2
- if QA criteria met = 1.

We also evaluated co-occurrence of attributes and state variables in available datasets. This step helped to identify datasets likrly to be useful when developing the attributes and establishing thresholds. These processes require complex information regarding performance as attributes and as estuarine health indicators.

2.6 Summarising information for recommendations

Responses to the four questions included in the technical survey (below) were summarised, by calculating the proportion (%) of respondents who answered "Yes" for each proposed attribute and state variable:

- Q6. Do you consider that a single robust method is used around New Zealand?
- Q7. For attribute variables - do you think it would be easy to predict this variable from upstream measures?
- Q8. Do you believe that natural long-term temporal patterns in this variable are understood or predictable?
- Q9. Do you believe that natural spatial patterns (for example estuary type, north vs south, east coast vs west coast) in this variable are understood or predictable?

These continuous values were converted to ranks using the following scale:

1. >0.8
2. 0.6 – 0.8
3. 0.5 – 0.6
4. 0.3 – 0.5
5. 0.2 – 0.3
6. <0.2

By considering the results of the ranking exercise, as well as expert opinion (including responses provided in the comments fields of the online survey and methodological factsheets), we are able to provide recommendations in the last section of the report on:

- Short-term (within the project lifetime) and long-term (beyond the project) strategies for filling the key gaps and addressing identified bottlenecks.
- Monitoring guidelines for attributes, including caveats and further development requirements.

3 Results

3.1 Review of methods used for monitoring of attributes

An overview of methods, as well as information regarding the consistency, potential bottlenecks and expert considerations for improved monitoring of attributes is provided in 10 detailed factsheets (Appendix A). Additional information regarding methods used to monitor state variables is summarised in a table (Appendix B).

We conclude that considerable lack of consistency in sampling and analytical methods and computational approaches exist for most attributes and state variables. In most instances, standardised guidelines do not exist (e.g., for FIB indicators, macroalgae, macrofauna, sediment chlorophyll a). Frequently, variations in sampling design, site selection, analytical resolution, and in reported measurement units impede comparison of acquired data.

Identified methodological bottlenecks are summarised in Table 3-1.

For most state variables not considered as attributes, consistent sampling and analytical methods exist, except for:

- dissolved oxygen
- redox potential discontinuity depth
- finfish diversity and abundance
- frequency of bathing beach closures
- frequency of customary harvest closures
- harvest area accessibility.

Following this review, we conclude that the main issue is inconsistent use of standardised sampling procedures and analytical protocols – this adversely affects our ability to compare data.

Table 3-1: Summary of potential bottlenecks, caveats and recommendations for prioritised attributes. Possible fixes (within or beyond the Stage 2 of the project) are provided in Section 4 of this report.

<table>
<tr><th rowspan="2">Attributes</th><th colspan="4">Consistent and/or accredited methods exist in NZ
Yes/No/Not Applicable</th><th rowspan="2">Potential bottlenecks</th><th rowspan="2">Caveats and recommendations</th></tr>
<tr><th>Sampling design</th><th>Sampling</th><th>Lab analyses</th><th>Computation</th></tr>
<tr><td>Water nutrients</td><td>N</td><td>N</td><td>Y</td><td>NA</td><td rowspan="4">Respond to upstream loading, but depend on intrinsic estuarine processes and inputs from other sources.
Difficult to separate its response to different stressors (e.g., sediment vs. nutrient loading).
High temporal-spatial variability.
Variable analytical methods with limited comparability.
Inconsistent sampling methodology in combination with high spatiotemporal variability limits comparability of datasets.
Water column nutrient concentrations do not necessarily reflect the quantity of nutrients available to primary producers.
Spot sampling for TSS is likely to be negatively impacted by the high temporal variation in suspended sediment loads.</td><td rowspan="4">Standardisation of sampling and analysis required.
Setting of thresholds should account for hydrosystem type.
Continuous measures are preferable over spot sampling.
Better understanding of the relationship to values and stressors is needed.
Existing datasets might not be suitable for threshold setting nationally.</td></tr>
<tr><td>Water Chl-a</td><td>N</td><td>N</td><td>Y</td><td>NA</td></tr>
<tr><td>Water clarity</td><td>N</td><td>Y</td><td>NA</td><td>NA</td></tr>
<tr><td>TSS</td><td>N</td><td>N</td><td>NA</td><td>NA</td></tr>
<tr><td>Water FIB</td><td>Y</td><td>Y</td><td>Y</td><td>Y</td><td>Affected by surrounding catchments and land use, but also characteristics of an estuary.
Lack of consensus around the state measure, statistic and minimum sample size to report.
High temporal-spatial variability.
Time-consuming analytical approaches impede timely warnings.</td><td>Modelled FIB estimates can be used for developing the attribute, measured concentrations used for state assessment.
Using qPCR or real-time in-situ sensors for rapid results and/or remote monitoring.
Molecular markers for source tracing and direct measurements of pathogens vs. indicators of pathogen presence.</td></tr>
</table>

Attributes	Consistent and/or accredited methods exist in NZ Yes/No/Not Applicable				Potential bottlenecks	Caveats and recommendations
	Sampling design	Sampling	Lab analyses	Computation		
					Contamination of samples during laboratory incubation possible, resulting in false alarms.	
Macroalgae	Y	Y	NA	Y	- Respond to eutrophication (nutrient loads), but expression of the response can be affected by hydrosystem characteristics (type, physical, hydrodynamic conditions, etc.). - Variation in the application of the sampling design. - Lack of adequate training for consistent assessment of macroalgae. - Absence of a collated national dataset of existing data, uncertainty or inconsistency in the ground truthing undertaken in different estuaries.	- Development of standardized methods for the field measurement of biomass, percentage cover. - Development of integrated GIS based mapping outputs and calculators. - Improve understanding of the relationship between nutrient loads and ecological response (including macroalgal growth). - Thorough assessment of ecological threshold responses over all estuary types. - Use of drones and/or remote sensing tools.
Macrofauna	Y	Y	N	N	- Integrate complex environmental conditions and represent benthic health, but may be difficult to distinguish stressor-specific response. - Spatial variability. - Lack of consistency in sampling design and taxonomic resolution.	- Comprehensive and consistent testing of existing macrofauna metrics for developing robust attributes/ state variables. - Better understanding of the broad-scale spatial variability and responses to stressors. - Suggested eutrophication-related thresholds need to be calibrated for different stressors/ estuary type/ bioregions. - Use of molecular ID methods for improved (and consistent) resolution.
Sediment Chl-a	Y	Y	N	N	- Responds to nutrient load, but can be affected by other stressors (upstream and/or estuarine). - High natural variability (temporal/spatial). - Lack of consistency in sampling design (esp. spatial extent, frequency, timing) and sample processing. - Inconsistent units used in reports.	- Standardise sampling protocols, analyses and reported units. - Sampling of subtidal sites should be considered. - Develop national thresholds, calibrated for different stressors/ estuary types/ bioregions.

Attributes	Consistent and/or accredited methods exist in NZ Yes/No/Not Applicable				Potential bottlenecks	Caveats and recommendations
	Sampling design	Sampling	Lab analyses	Computation		
Mud content/ grain size	Y	Y	N	Y	- Indicative of habitat change and sediment supply, can be used as a surrogate for sediment accumulation, but can be effected by *in situ* processes and stressors. - High natural within-year and between-year variability without a strong predictable pattern. - Lack of consistency in sampling design (especially frequency and site selection) and analytical methods (results are not necessarily comparable). - Inconsistent size fractions are used for reporting "mud", make comparison of results difficult.	- Standardise analytical method. - Better understanding of the effect of temporal frequency and replication. - Sampling subtidal sites should be considered for certain estuaries. - Existing thresholds (ETI) may need to be calibrated for different stressors/ estuary type/ bioregion.
Sediment deposition rate (including Annual Average Sedimentation Rate, AASR)	N	N	N	N	- Responds to land disturbance in the catchment and associated sediment loads, however can be insufficient as a standalone measure for managing the stressor, other sediment stress related elements need to be considered. - Lack of nationally consistent data for land cover, different models used to estimate sediment loads. - Poor, inconsistent and often incomparable measurement techniques - Metrics unrelated to specific estuary conditions.	- Standardise measurement method. - Use of multiple complementary methods (fine-scale and broad-scale) to increase confidence and account for spatial variability. - A future-focused long-term monitoring to establish meaningful trends. - Refinement of existing models to reduce uncertainty and increase accuracy of predictions. - Collation of national data to enable refinement of proposed thresholds (ETI) for management.
Sediment metals	Y	Y	N	N	- Responds to human-induced changes in land-use and land-derived contamination, however other sources of contamination (e.g., stormwater) may make discernment of upstream effects difficult. - Lack of consistence in sampling (especially subtidal). - Inconsistent analytical methods.	- Standardise sampling and analyses. - Investigate the influence of the analysis of different grain size fractions on the results. - Validate national trigger values (ANZECC 2000) for threshold setting (ecological effect can occur at lower metal values than indicated).

Attributes	Consistent and/or accredited methods exist in NZ Yes/No/Not Applicable				Potential bottlenecks	Caveats and recommendations
	Sampling design	Sampling	Lab analyses	Computation		
Shellfish metals and other contaminants	N	N	N	N	- Respond to human-induced changes in land-use and land-derived contamination, but can be affected by other internal and external sources of contamination. - Limited data available for establishing thresholds and for detection of trends. - Loss of caged mussels due to vandalism or extreme weather. - Limited number of contaminants with human health standards. - Some chemical contaminants of potential concern may bioaccumulate to limited extent – difficult to detect. - Varying detection limits between laboratories.	- Implementation of a standardised long-term monitoring programme. - Use of both resident and caged shellfish for biomonitoring. - Include species with different feeding mode to distinguish water and sediment pathways for contaminant exposure. - Selection of appropriate reference sites.
Shellfish FIB	N	Y	Y	Y	- Samples are often biased toward good weather conditions and when shellfish are being harvested. - Potential false alarms due to contamination issues during processing. - Time-integrated bioaccumulation makes it difficult to distinguish effect of a particular stressor. - Differential depuration between indicators and pathogens breaks down relationship between traditional bacterial indicators and viral pathogens. - Time-consuming laboratory analyses (culture-based) do not allow "real time" response. - Lack of consistency in sampling design (spatial extent, number of replicates, timing).	- Standardised (sufficient and representative) sampling design. - A faster FIB assessment method, application of molecular techniques. - Modelled estimates for shellfish FIB concentrations in response to upstream loading. - Consider supporting environmental information (e.g., wave action, climate, tidal state).

3.2 Existing datasets for prioritised attributes and state variables

Using available information, we reviewed currently available data to determine overall availability and temporal coverage. We also analysed regional and typological coverage represented in these data for each prioritised attribute/ state variable. The data considered here are not exhaustive, because data collection was focused primarily on the previously identified regional/typology gaps (see Section 2.2.2). Further targeted effort to acquire additional data was anticipated in Stage 2.

Overall, we identified 41 groups of data with potential for use in Stage 2. These include previously identified, "ready-to-use" datasets and individual data files/reports acquired through targeted collection. In terms of availability, 18 data classes were assigned to Category 1, 8 were assigned to Category 2, and 15 were included in Category 3 (Table 3-2 and Table 3-5, Figure 3-1).

Table 3-2: Summary of identified data categories for attributes and state variables. *Potential attributes for future development.

Potential use	Variables prioritized in Stage 1A	Category(s) [10] of identified datasets
Attributes	Water nutrients (TN, TP)	1, 2, 3
	Water Chl-a	1, 3
	Water clarity	1, 2, 3
	Total suspended solids	1, 3
	Water faecal indicator bacteria (FIB)	1, 2, 3
	Macroalgae (e.g., Ulva)	1, 3
	Macrofauna	1, 3
	Sediment Chl-a	2
	Mud content/grain size	1, 3
	Sediment deposition rate (incl. AASR)	1, 3
	Sediment metals	1, 2, 3
	Shellfish metals	1, 3
	Shellfish faecal indicator bacteria (FIB)	1, 2, 3
	*Emerging contaminants**	2
	*Molecular FIB markers**	1
State variables	Dissolved oxygen	1, 3
	Sediment nutrients (TN, TP)	1, 3
	Sediment TOC	1, 3
	Sediment sulphides	2, 3
	Redox potential discontinuity depth	1, 2, 3
	Extent of dominant substrate types (e.g., mud)	2, 3
	Extent of habitats (e.g., seagrass beds)	1, 2, 3
	Finfish diversity and abundance	2, 3
	Shellfish distribution and abundance	3
	Frequency of bathing beach closures	1
	Frequency of customary harvest closures	-
	Harvest area accessibility	-

[10] Category 1: *data are useful, available and collected (was provided for the project use) along with accompanying metadata*;
Category 2: *data are potentially useful and existing, can be made available later*;
Category 3: *data exist but availablility timeline and / or usefulness is uncertain*

The Category 1 data represented most (19) of the 27 prioritised attributes/ state variables. The attributes/ state variables that were not well represented and which were included in Category 2 or Category 3 included: finfish diversity and abundance (Category 2 and 3), extent of dominant substrate types (2 and 3), emerging contaminants (2), sediment chl-a (2), sediment sulphides (2 and 3) and shellfish distribution and abundance (3). Attributes/state variables for which no data were identified included: shellfish molecular markers, frequency of harvest closures, and harvest area and accessibility.

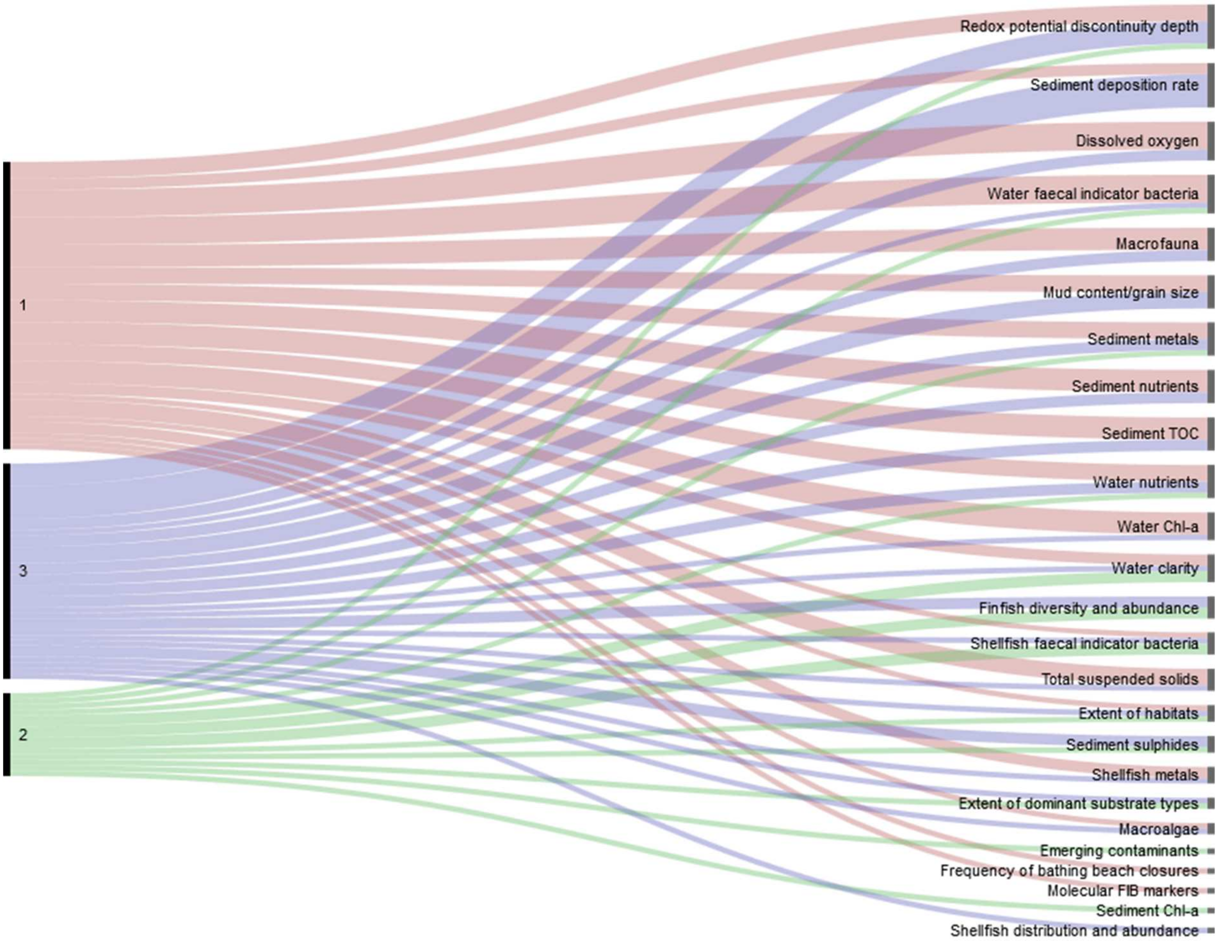


Figure 3-1: Alluvial diagram showing coverage of attributes and state variables (right) by different data categories (left). Data categories were assigned as follows: 1 - useful, available and collected; 2 – potentially useful, data can be available later; 3 – data exists but availability timeline / usefulness is uncertain. The width of the ribbons is representative of the number of datasets/individual data files identified.

Overall, the datasets spanned the thirty-year period from 1987–2017 (Table 3-3). Of the Category 1 and 2 datasets for which the temporal extent was known, approximately one third (35%) were collected over relatively short (≤ one year) timeframes, while another third (35%) were collected over relatively long timeframes (≥ 10 years).

Although category 1 data were obtained for all 16 regions, and Category 2 data for nearly all (14), not all attributes/state variables were represented in each region. Regions (represented by Councils) for which least data existed were GDC and TRC.

All estuary types were represented in Category 1 and 2 data, with data for SIDE estuaries the most abundant, available for 24 of 27 attributes/ state variables, exceptions being: frequency of harvest closures, shellfish distribution and abundance, and harvest area and accessibility (Table 3-3).

From the gaps overview, we identified critical data shortages for the following attributes:

- *shellfish metals*
- *shellfish FIB*
- *emerging contaminants*

and state variables:

- *sediment sulphides*
- *shellfish distribution and abundance*
- *frequency of harvest closures*
- *harvest area accessibility.*

Combined information (supporting indicators) did not exist for many variables (Table 3-4). For example, data for the attributes "shellfish metals" and "shellfish FIBs" are not accompanied by any information for estuary health indicators. "Sediment sulphides" and "Extent of dominant substrates" are also reported as stand-alone variables in the available datasets. Absence of combinatory information is considered an impediment to further development of selected attributes/ state variables using the available datasets, because absence of data prevents relating fine-scale attributes to estuary condition and/or habitat type.

Table 3-3: Availability of data on attributes and state variables for different types of estuaries and regional councils. Y – data available, N - data not available. Data categories 1 and 2 were considered here, as insufficient information for category 3 data did not allow comprehensive assessment of regional/typology coverage. It should be noted though that some of the gaps presented here could be covered in category 3 data.

Potential use	Attribute/ State variables	Data by estuary type			Data by regions															
		SSRTRE	SIDE	DSDE	NRC	AC	WRC	BOPRC	TRC	HRC	GDC	HBRC	GWRC	MDC	NCC	TDC	ECAN	WCRC	ORC	ES
Attributes	Water nutrients (TN, TP)	Y	Y	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	N	N	Y	Y	Y	N
	Water Chl-a	Y	Y	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Water clarity	Y	Y	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Total suspended solids	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	N	N	Y	Y	Y	N
	Water faecal indicator bacteria	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Macroalgae (e.g., Ulva)	Y	Y	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Macrofauna	Y	Y	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Sediment Chl-a	Y	Y	Y	Y	Y	N	Y	N	N	N	Y	N	Y	Y	Y	Y	Y	Y	Y
	Mud content/grain size	Y	Y	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Deposition rate (incl. AASR)	Y	Y	Y	N	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Sediment metals	Y	Y	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Shellfish metals	N	Y	N	N	N	N	Y	N	N	N	N	N	N	N	N	Y	N	N	N
	Shellfish faecal indicator bacteria	Y	Y	N	N	Y	N	Y	N	N	N	N	N	N	N	N	N	N	N	N
	Emerging contaminants	N	Y	N	Y	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
	Molecular FIB	Y	Y	Y	Y	Y	Y	N	Y	N	N	Y	Y	Y	N	N	Y	Y	N	Y

Potential use	Attribute/ State variables	Data by estuary type			Data by regions															
		SSRTRE	SIDE	DSDE	NRC	AC	WRC	BOPRC	TRC	HRC	GDC	HBRC	GWRC	MDC	NCC	TDC	ECAN	WCRC	ORC	ES
	markers (water)																			
State variables	Dissolved oxygen	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Sediment nutrients (TN, TP)	Y	Y	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Sediment TOC	Y	Y	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Sediment sulphides	N	Y	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N	N	N
	Redox potential discontinuity depth	Y	Y	Y	Y	Y	N	Y	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Extent of dominant substrate	Y	Y	Y	N	N	N	N	N	Y	N	N	Y	N	N	Y	N	Y	N	Y
	Extent of habitats (e.g., seagrass beds)[11]	Y	Y	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Finfish diversity and abundance	Y	Y	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Shellfish distribution and abundance	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
	Frequency of bathing beach closures	Y	Y	Y	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Frequency of harvest closures	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
	Harvest area accessibility	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N

[11] Based mostly on the ETI Tool2 Input demonstrational data (macroalgae cover only). The QA of this data is uncertain. QA-compliant data is assumed from ORC, ECAN, MDC and NCC (Wriggle broad-scale), but not available at the moment.

Table 3-4: Co-occurrence matrix of attributes and state variables in the identified available datasets (categories 1 and 2) that were considered for future use. Figures in the cells indicate the number of datasets in which the variables co-occur. Cells within the red outlined area represent co-occurrence of attributes and state variables.

Variables Prioritised in Stage 1		Water nutrients (TN, TP)	Water Chl-a	Water clarity	Total suspended solids	Water faecal indicator bateria (FIB)	Macroalgae (e.g. Ulva)	Macrofauna	Sediment Chl-a	Mud content/grain size	Deposition rate (incl. AASR)	Sediment metals	Shellfish metals	Shellfish faecal indicator bacteria (FIB)	Dissolved oxygen	Sediment nutrients (TN, TP)	Sediment TOC	Sediment sulphides	Redox potential discontinuity depth	Extent of dominant substrate types (e.g. mud)	Extent of habitats (e.g. seagrass beds)	Finfish diversity and abundance	Frequency of bathing beach closures
Attributes	Water nutrients (TN, TP)																						
	Water Chl-a	2																					
	Water clarity	1	1																				
	Total suspended solids	3	2	1																			
	Water faecal indicator bateria (FIB)	3	1	1	2																		
	Macroalgae (e.g. Ulva)	0	1	0	0	0																	
	Macrofauna	0	0	0	0	0	0																
	Sediment Chl-a	0	0	0	0	0	0	0															
	Mud content/grain size	0	0	0	0	0	0	2	0														
	Deposition rate (incl. AASR)	0	1	0	0	0	1	0	0	0													
	Sediment metals	0	0	0	0	0	0	2	0	3	0												
	Shellfish metals	0	0	0	0	0	0	0	0	0	0	0											
	Shellfish faecal indicator bacteria (FIB)	0	0	0	0	0	0	0	0	0	0	0	0										
State variables	Dissolved oxygen	3	4	1	3	2	1	0	0	0	1	0	0	0									
	Sediment nutrients (TN, TP)	0	1	0	0	0	1	2	0	3	1	3	0	0	1								
	Sediment TOC	0	1	0	0	0	1	2	0	3	1	3	0	0	1	4							
	Sediment sulphides	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0						
	Redox potential discontinuity depth	0	1	0	0	0	1	2	1	1	1	1	0	0	1	2	2	0					
	Extent of dominant substrate types (e.g. mud)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
	Extent of habitats (e.g. seagrass beds)	0	1	0	0	0	1	0	0	0	1	0	0	0	1	1	1	0	1	1			
	Finfish diversity and abundance	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
	Frequency of bathing beach closures	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

Most (25) of the data considered fulfilled the QA criteria (i.e., overall QA rating was 'Yes'), indicating likely usefulness for Stage 2 of the project. The remaining data were given a QA rating of TBC (uncertain), pending provision of further information.

An overview of the temporal extent, availability and potential importance of data, as well as a ranking of results is presented in Table 3-5. Data were ranked in terms of potential importance by applying the sum of the following scores:

- If number of estuaries covered >1 = 1.
- If temporal data (> 5 years) =1.
- Types of estuaries covered:
  - 2 types = 1
  - 3 types = 2
- Number of regions covered:

    - 2 regions = 1
    - > 2 regions = 2
- Number of variables covered:
    - 2 variables = 1
    - >2 variables = 2
- If QA criteria met = 1

The overall importance is considered greater for data where higher ranks were assigned.

Table 3-5: Overview of the temporal extent, availability and potential importance of the identified data, based on acquired information (data or metadata). Note these timeframes represent the data/project as a whole and do not account for variation in the temporal extent of individual attributes / state variables within a dataset. * = years identified from metadata only. TBC indicates that information was not provided (or was insufficient) and needs to be confirmed.

Dataset name (as per metadata table submitted to MfE)	Availability category	Overall temporal extent of dataset[12]	Comments on availability and potential importance (typology/regional coverage, variables represented, QA results)	Importance rank
MfE WQ dataset (compilation of Council data)	1	1973 - 2016	Compiled long-term dataset, multiple variables, multiple estuaries, different types/regions covered. Data available, QA criteria met.	9
Mussels and oysters 1990-2016 comparison project (Tauranga, Waikareao)	1	1990 and 2016	Project data from 2 estuaries, single variable (shellfish metals)/ type/ region represented. Data available, QA criteria met.	2
BOPRC shellfish bacteria	1	1994-2015	Data from several estuaries (SIDE and SSRTRE), same region, single variable. Data available, QA criteria met.	4
BOPRC harbour clarity	1	1999-2015	Data from several estuaries (SIDE and SSRTRE), same region, single variable. Data available, QA criteria met.	4
OTOT (compilation of Council data)	1	2001 - 2016	Compiled long-term dataset, multiple variables, multiple estuaries, different types/regions covered. Data available, QA criteria met.	9
Estuary Project - water quality (Catlins, Tokomairiro, Taieri, Kaikorai, Waikouaiti, Shag, Kakanui)	1	2004-2009	Compiled dataset from 1 region, multiple variables, several estuaries (one off sampling in each estuary), different types covered. Data available, QA criteria met.	6
Estuary Project - field data (Catlins, Tokomairiro, Taieri, Kaikorai, Waikouaiti, Shag, Kakanui)	1	2004-2009	Data from 1 region, 2 variables, several estuaries (one off sampling in each estuary), different types covered. Data available, QA criteria met.	5

[12] Within large datasets, temporal extent of data for particular variables might not correspond to the overall extent. In such cases, variable- specific temporal coverage was considered for ranking the importance, a corresponding comment was added where applicable.

Dataset name (as per metadata table submitted to MfE)	Availability category	Overall temporal extent of dataset[12]	Comments on availability and potential importance (typology/regional coverage, variables represented, QA results)	Importance rank
Greater Wellington recreational water quality for bathing sites	1	2006-2017	Compiled dataset from 1 region, 1 variable, several estuaries (SIDE and DSDE). Data available, QA criteria met.	4
Lyttelton Harbour - bed level monitoring	3	2007*	Short time scale bed level data (weeks to months), 1 variable/estuary. QA uncertain (not enough information).	0
ENV01-08002- shellfish data (Avon-Heathcote Estuary)	1	2008, 2010, 2012, 2014	Project data from 1 estuary (SIDE), single variable (shellfish metals. Data available, QA criteria met.	1
Sediment deposition data (Auckland, Waikato, Tasman)	1	2008-2014 (shorter for some estuaries)	Data from several estuaries/types/2 regions, single variable (not temporal – different years in different estuaries). Data available, QA criteria met.	5
MST Tools project	1	2011	Project data from several estuaries/types/regions, single variable (molecular markers). Data available, QA criteria met.	6
Porirua harbour SOE	1	2011-2013	Compiled dataset from 1 estuary (SIDE), several variables. Data available, QA criteria met.	3
Marine bathing beach data	1	2013-2016	Compiled dataset from multiple estuaries/regions/types, 2 variables. Data available, QA criteria met.	7
BOP Estuarine benthic ecosystem monitoring	1	2013-2017	Compiled dataset from 3 estuaries (SIDE and SSRTRE), 1 region, multiple variables. Data available, QA criteria met. Some double up with OTOT dataset.	5
Mahakipawa Estuary	1	2017	Data from 1 estuary (data for other MDC estuaries may be extracted from online reports), 1 variable. Data available, QA criteria met.	1
Porirua Harbour microbial results	1	2017	Data from 1 estuary, 1 variable. Data available, QA criteria met.	1
Wriggle - Manawatu River Estuary	1	2017	Data from 1 estuary, several variables. Data available, QA criteria met. Double up with Wriggle Fine scale data.	3
ETI Tool2 input data	1	TBC	Dataset from multiple estuaries/types/regions, multiple variables. Demonstration data only available, QA uncertain (not enough information in the demonstration data), likely some double up with other data.	7

Dataset name (as per metadata table submitted to MfE)	Availability category	Overall temporal extent of dataset[12]	Comments on availability and potential importance (typology/regional coverage, variables represented, QA results)	Importance rank
Auckland Shellfish and Sediment Contaminant Monitoring Programme	2	1987-2011*	Data from 3 estuaries (SIDE), 1 region, 3 variables (shellfish and sediment metals, emerging contaminants). Likely available from Marcus Cameron later this year, QA criteria met (preliminary assessment).	5
NIWA estuarine intertidal fish survey	2	2001-2007*	Dataset from multiple estuaries/types/regions, 2 variables. Available from Malcolm Francis from October 2017, QA criteria met (preliminary assessment).	8
NIWA juvenile rig survey	2	2011*	Dataset from multiple estuaries (SIDE and DSDE)/regions, 2 variables. Available from Malcolm Francis from October 2017, QA criteria met (preliminary assessment).	7
Porirua Harbour subtidal survey	2	2015*	Data from 1 estuary (SIDE), potentially also data from Wellington harbor available, 1 variables (sediment sulphides). Potentially available later through Claire Conwell, QA uncertain (not enough information).	0
Horizons broad scale data	2	2016*	Data from several estuaries/types, 1 region, 2 variables (extent of habitats and extent of dominant substrates). Potentially available later this year. Potential double up with Wriggle Broad scale data, QA criteria met (preliminary assessment).	5
Marlborough - LAWA Coastal module	2	TBC	Data from 2 estuaries (DSDE), 1 region, 2 variables. Data may be extracted from online, QA uncertain (not enough information).	2
Councils data Chl-a and RPD data (individual reports)	2	TBC	Data from multiple estuaries/types/regions, 2 variables. Data available but need to be compiled, QA criteria met (preliminary assessment).	7
CPUE analyses for commercial fish (e.g., flounder)	3	1990-2014*	Data from 1 estuary. Not known at the time whether data files in the reports are available and able to be used for this project. QA criteria met (preliminary assessment).	2
Grey mullet survey across the North Island and upper South Island	3	2010*	Data from several estuaries/types/regions, 1 variable (finfish diversity). Not known at the time whether data files in the reports are available and able to be used for this project, QA criteria met (preliminary assessment).	6
Shellfish metals Environment Southland	3	2013*	Data from 1 region (number of estuaries/types – to be confirmed). Data not received after data request (availability timeline uncertain), QA criteria met (preliminary assessment).	1
Sediment sulphides Southland	3	2016*	Data from 1 region, 5 estuaries (types – to be confirmed). Data not received after data request, QA uncertain (not enough information).	1

Dataset name (as per metadata table submitted to MfE)	Availability category	Overall temporal extent of dataset[12]	Comments on availability and potential importance (typology/regional coverage, variables represented, QA results)	Importance rank
Environment Southland - Sedimentation rates (plates and historic)	3	TBC (>10 years duration)*	Data from several estuaries (SIDE and SSRTRE), multiple variables. Data not available at present, some overlap with Wriggle data. QA uncertain (not enough information).	5
NIWA North Island - sedimentation - coring data	3	TBC (historic data)	Data from multiple estuaries/types, 1 region, 1 variable. Data not available at present. QA uncertain (not enough information).	4
Greater Wellington Council - sediment accumulation plates	3	TBC (5 to 7 years duration for different estuaries)*	Data from 4 estuaries/2 types, 2 variables (sediment deposition rate and sediment metals). Data not available at present, some overlap with Wriggle data. QA uncertain (not enough information).	4
Wriggle Fine scale data	3	TBC	Data from multiple estuaries/types/regions, multiple variables. Not available at present but could be in the future with additional work (need to be compiled). QA criteria met (preliminary assessment). Some double up with other datasets e.g., OTOT.	8
Wriggle Broad scale data	3	TBC	Data from multiple estuaries/types/regions, multiple variables. Not available at present but could be in the future with additional work (need to be compiled). QA criteria met (preliminary assessment). Some double up with other datasets e.g., OTOT.	8
Wriggle OMBT data	3	TBC	Data from multiple estuaries/types/regions, multiple variables. Not available at present but could be in the future with additional work (need to be compiled). QA criteria met (preliminary assessment). Some double up with other datasets e.g., OTOT.	8
Water clarity data New River Estuary (TSS, water TN and TP, DO)	3	TBC	Data from 1 estuary (SIDE), several variables. Data availability timeline is uncertain, may be able to get permission from Invercargill City Council in the future. QA uncertain (not enough information).	2
NIWA historic sedimentation (Lyttelton Harbour)	3	TBC (over last 400 years)*	Single sampling (core analysis), 1 estuary/variable. Data availability timeline and QA are uncertain.	1
MPI shellfish information	3	TBC (2012-2017)*	Multiple surveys with different timeframes (there might also be some data prior to 2012), 1 variable (shellfish distribution and abundance). Not known at the time whether data files in the reports were available and able to be used for this project. QA uncertain (not enough information).	0

Dataset name (as per metadata table submitted to MfE)	Availability category	Overall temporal extent of dataset[12]	Comments on availability and potential importance (typology/regional coverage, variables represented, QA results)	Importance rank
Ecology of faecal indicators in estuarine waters and shellfish (University of Otago- Master's thesis)	3	TBC	Two variables (water and shellfish FIB), not enough information on availability and usefulness for the project.	1
Marlborough shellfish FIB	3	TBC	Data not received after data request, potential double up with the MfE dataset. Not enough information on availability and usefulness for the project.	0

Based on the results above, the most promising data for baseline assessment and for further development of attributes and state variables in Stage 2 of the project are summarised in Table 3-6:

Table 3-6: Most promising data identified for use in Stage 2 of the project. Variables that scored >5 in Table 3-5 were selected.

Variables		Suggested useful datasets	Comments
Attributes	Water nutrients (TN, TP)	MfE WQ dataset, Wriggle Fine scale data, Estuary Project - water quality	Additional effort might be needed to compile Wriggle data into a single dataset
	Water Chl-a	MfE WQ dataset, Wriggle Fine scale data, ETI Tool2 input data	Additional effort might be needed to compile Wriggle data into a single dataset; QA for ETI data to be confirmed
	Water clarity	MfE WQ dataset, Wriggle Fine scale data	Additional effort might be needed to compile Wriggle data into a single dataset
	Total suspended solids	MfE WQ dataset, Estuary Project - water quality	
	Water faecal indicator bacteria (FIB)	MfE WQ dataset, Estuary Project - water quality, Marine bathing beach data	
	Macroalgae (e.g., Ulva)	Wriggle OMBT data	Additional effort might be needed to compile Wriggle data into a single dataset
	Macrofauna	OTOT, Wriggle Fine scale data	Taxonomic lumping will likely be required for OTOT; Additional effort might be needed to compile Wriggle data into a single dataset
	Sediment Chl-a	Council Sediment Chl-a data, ETI Tool2 input data	Effort needed to tease apart different methods and compile the datasets from individual reports; QA for ETI data to be confirmed, potential double-up with councils' data
	Mud content/grain size	OTOT, Wriggle Fine scale data	Inconsistent lab analyses in OTOT, in some instances can be incomparable; Additional effort might be needed to compile Wriggle data into a single dataset
	Deposition rate (incl. AASR)	Wriggle Fine scale data	Additional effort might be needed to compile Wriggle data into a single dataset
	Sediment metals	OTOT, Wriggle Fine scale data	Inconsistent lab analyses in OTOT, in some instances can be incomparable; Additional effort might be needed to compile Wriggle data into a single dataset
	Molecular FIB markers (water)	MST Tools project	
State variables	Dissolved oxygen	MfE WQ dataset, Wriggle Fine scale data, Estuary Project - water quality	Additional effort might be needed to compile Wriggle data into a single dataset
	Sediment nutrients (TN, TP)	OTOT, Wriggle Fine scale data, ETI Tool2 input data	Two types of nitrogen reported in OTOT, in some instances can be incomparable; Additional effort might be needed to compile Wriggle data into a single dataset; QA for ETI data to be confirmed, potential double-up with OTOT and Wriggle data
	Sediment TOC	Wriggle Fine scale data, ETI Tool2 input data	Additional effort might be needed to compile Wriggle data into a single dataset;

Variables		Suggested useful datasets	Comments
			QA for ETI data to be confirmed, potential double-up with Wriggle data
	Sediment sulphides	Wriggle Fine scale data	Additional effort might be needed to compile Wriggle data into a single dataset
	Redox potential discontinuity depth	Wriggle Broad scale data, Wriggle OMBT data, Wriggle Fine scale data, Council RDP data, ETI Tool2 input data	Additional effort might be needed to compile Wriggle data into a single dataset; Effort needed to compile the council datasets from individual reports; QA for ETI data to be confirmed, potential double-up with councils' and Wriggle data
	Extent of dominant substrate types (e.g., mud)	Wriggle Broad scale data	Additional effort might be needed to compile Wriggle data into a single dataset
	Extent of habitats (e.g., seagrass beds)	Wriggle Broad scale data	Additional effort might be needed to compile Wriggle data into a single dataset
	Finfish diversity and abundance	NIWA nationwide estuarine intertidal fish survey, NIWA juvenile rig survey, Grey mullet survey across the North Island and upper South Island	Additional effort might be needed to compile Grey mullet data

However, prior to use in Stage 2 and Stage 3, the ecological relevance and representativeness of these data should be considered further. For instance, most of the EMP data (OTOT dataset) have been collected from the dominant habitat in each estuary. These may not necessarily be the most susceptible to upstream inputs; there may also be an interaction between susceptibility and overall estuarine condition. Potential issues in the datasets intended for further use would be considered in greater detail during the Stage 2 analysis.

3.3 Expert survey

In total, 19 survey responses were received (11 from the project partners and 8 from external technical experts, see Appendix C for details). We acknowledge that answers provided to specific survey questions could be driven by personal interpretation of scale and understanding of estuary processes. Therefore, responses were considered as an overall litmus test – 'Yes' *versus* 'No', without emphasizing or considering all the possible nuances between those extreme ends.

Analysis of the results from Part A of the survey (see Appendix D) was aimed at validating the prioritisation of variables conducted in the first phase of Stage 1. This indicated that all the attributes considered were viewed as suitable by experts (all of them were also considered suitable as state variables). However, comments provided were considered further, and this influenced selection as well. Examples include:

- *Some of the attributes are correlated with each other (i.e., not independent), e.g., water clarity and TSS, similarly RPD and sulphides, mud and TOC.*
- *"Presence of macroalgae" may need to be clarified as opportunistic or bloom-forming macroalgae, if these are to be considered as an indicators of health.*
- *Shellfish and sediment metals should be considered as 'toxicants' to allow flexibility for inclusion of additional existing toxicants, as well as new contaminants. Suitability of contaminants in shellfish as an indicator varies*

depending on contaminant and established guidelines for effects on shellfish and variable guidelines for human health.

All proposed state variables were considered suitable with the exception of "harvest area accessibility"[13], although there was less agreement regarding the suitability of "dissolved oxygen" and "finfish diversity and abundance" as state variables.

In Part B of the survey, the ability to predict attributes from the upstream measurements was considered rather low (with a slightly higher rate of "Yes" responses for mud content/ grain size and water nutrients (Figure 3-2)). It was commented however, that high-level predictions should be possible for most attributes, e.g., direction and relative degree of change, if a relevant upstream measure is available. Overall, likely accuracy of predictions of state within an estuary will be low or variable (because of high spatial and temporal variability), even though trends can be well predicted at national and local levels for many attributes.

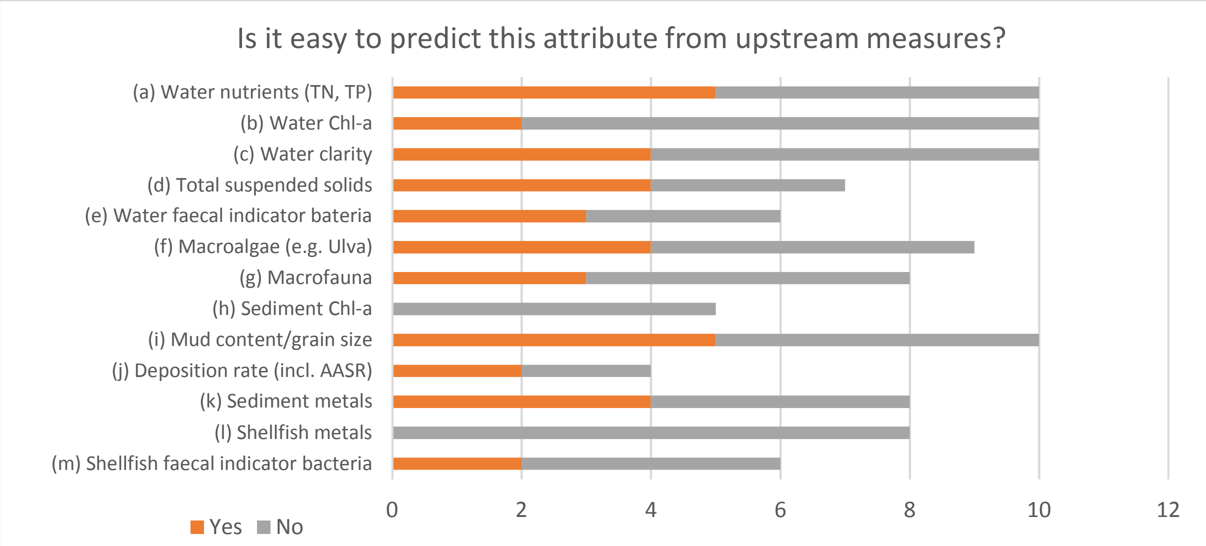


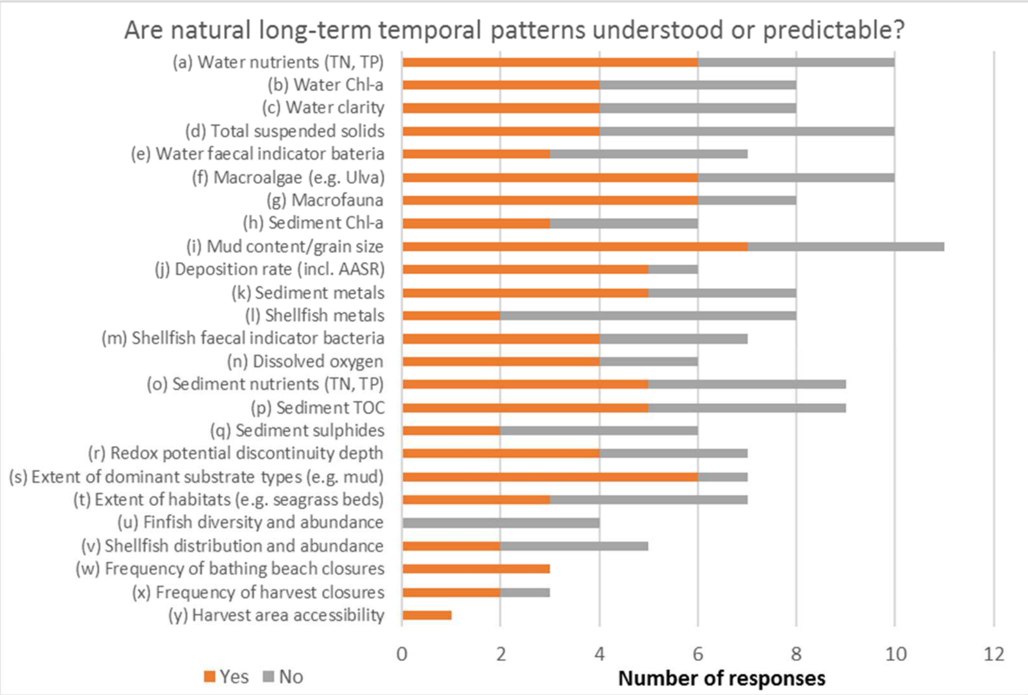
Figure 3-2: Online survey results: predictability of an attribute from upstream measures.

The opinion of respondents around temporal and spatial patterns (Figure 3-3) varied considerably, with highest uncertainty associated with spatial patterns of variables arising from natural processes. Variables with especially uncertain responses included:

Attributes:	State variables:
Water Chl-a	Sediment nutrients
Water clarity	Sediment TOC
Total suspended solids	Sediment sulphides
Water FIB	RPD
Macroalgae	Extent of habitats
Sediment Chl-a	Finfish diversity and abundance
Shellfish metals	Shellfish diversity and abundance

[13] Currently there are no existing information on this attribute, making its suitability assessment difficult to impossible.

This outcome could partially reflect variability of the attributes/ state variables and their response to particular pressures at a national level. Much greater certainty for specific attributes is expected at a local scale. The extent of habitats is simple to measure, but will be different in every estuary and will change over time in response to various stressors. This will result in high uncertainty at a national level, but much greater certainty for specific attributes at local scale.



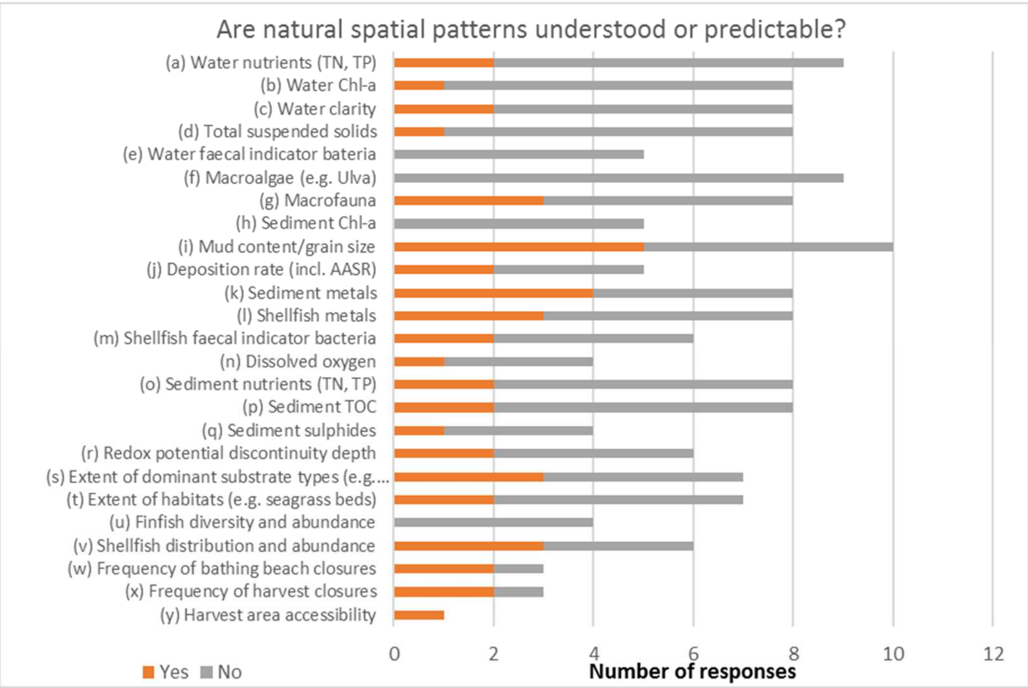


Figure 3-3: Online survey results: consensus of respondents around temporal and spatial patterns of considered attributes and state variables.

4 Considerations for Stage 2

The objective of the Stage 1 work was to identify potential and useful measures for attributes and/or state variables that would be developed further in Stage 2. These would only be considered useful if they directly linked to estuarine values and objectives, and for attributes, if the contaminants entering the estuary through freshwater inflows could be managed via limit setting to produce measurable changes in the attribute state[14]. Characteristics of good attributes include:

1. Ability to be predicted from upstream measures (albeit with some uncertainty) – applicable to attributes only.
2. Ability to be linked to one of the three values (ecosystem health, human health and mahinga kai) to the extent that the value could be predicted (albeit with some degree of uncertainty) - applicable to attributes and state variables.
3. Availability of a robust, standardised and cost-effective method - applicable to attributes and state variables.
4. Either low or predictable variability in space and time - predictable variability includes defining a specific sampling place within an estuary, a specific time of sampling (tidal or seasonal), or co-variables that can be used to explain variability, such as estuary typology - applicable to attributes and state variables.
5. Usefulness for establishment of management objectives, including setting numeric limits and thresholds - applicable to attributes only.

These characteristics were considered in the initial expert workshop, within the essential criteria used in the expert survey carried out and reported in Stage 1A (for details see Cornelisen et al. (2017)), and within expert survey in Stage 1B (this report).

4.1 Ability to predict from upstream measures

Proposed attributes were ranked according to the level of consensus technical experts exhibited in answering "yes" to the question that indicated whether candidate attributes were likely to be predictable in the estuary from upstream measures (Table 4-1). Technical experts generally either disagreed over this predictability (rank 3 and 4), or agreed that the variable was not predictable (ranks < 4). For both "sediment deposition rates" and "water column nutrients" this was partly because they can either be measured in the estuary or modelled directly from catchment and freshwater information. Water column nutrients are expected to be strongly affected by within-estuary dynamics - these would need to be measured in the upper estuary (close to the freshwater inflow) on the outgoing tide, limiting usefulness in estuaries with more than one significant freshwater inflow, and also making it difficult to validate modelled estimates. Annual average sedimentation rate (AASR) can currently be modelled across an estuary or for specific areas within estuaries, and these estimates have been validated in some estuaries, although this metric is strongly affected by estuary typology.

The variables most likely to benefit from Stage 2 analysis and collection of new data include:

[14] Limit here refers to upstream aspects to be managed and represents the maximum upstream loads which allow for freshwater and estuary objectives to be met.

- Sediment deposition rate
- Water nutrients (TN, TP)
- Total suspended solids
- Water faecal indicator bacteria
- Macroalgae
- Macrofauna
- Mud content/grain size (see Table 4-1).

Table 4-1: Consensus over whether proposed attributes were predictable from upstream measures. 1 = agreement that attribute is well predicted, 3 – 4 = disagreement about whether the attribute is well predicted, 6 = agreement that the attribute is not well predicted. For high rankings indicating potentially low predictability, a recommendation for improving the ranking is given, taking into account the results of methods overview and data gap analysis.

Potential attribute	Rank	Likely to be improved by Stage 2 analysis	Requires collection of new data	Requires new research	Use modelled information
Deposition rate (incl. AASR)	3		Yes – would need to train any future models with 'event' based data	Development of new method	Yes - but need to understand cost vs uncertainty implications
Water nutrients (TN, TP)	3	probably		Need to identify whether persulphate or TKN method most appropriate	Possibly, but this would require validation research
Total suspended solids	3	probably	Yes, targeted		Possibly, but requires targeted event based sampling to achieve peak flow volumes
Water faecal indicator bacteria	3	probably		Needs a review of the laboratory methods used to ensure consistency. Investigate relationships between pathogens and illness rate (recently funded MfE project)	Yes
Macroalgae	3	probably	Yes, targeted	Need training on measurements and ID to appropriate taxonomic resolution	
Macrofauna	3	probably	Yes, targeted		
Mud content/grain size	3	probably	Yes, targeted	Need a decision on a standard method	
Sediment metals	3	possibly	Yes, targeted		
Water clarity	4	probably	Yes, targeted	Need a decision on a standard method	
Shellfish faecal indicator bacteria	4	possibly	Yes, targeted		
Water Chl-a	5		Yes, targeted	Inter-laboratory QA	
Sediment Chl-a	6			Assess comparability of methods	
Shellfish metals	6		Yes, targeted		

4.2 Predictability of values from measures

Here we present a summary of the information gathered in the first workshop, evaluation of variables and survey (Stage 1A) from experts in estuarine ecosystem health, human health and mahinga kai (Table 4-2). A list of the participants is given in the first report (Cornelisen et al. (2017)).

Table 4-2: Values and the major stressors and variables that represent them. Bolded variables are candidate attributes from Report 1A (see Figure 1-1) and those not bolded include potential state variables and/or supplementary variables used in estuary monitoring. *For water nutrients, a proxy, such as modelled potential nutrient concentrations were suggested in Stage 1A report (Cornelisen et al. 2017), wherever direct spot measurements are advisable for water nutrients as state attribute.

Value/objective	Major stressor	Most direct measure	Other measures
Ecosystem health			
Good water quality (clear and uncontaminated)	Nutrients, Sediments	**Water clarity**	**Water Chl-a, TSS, Water nutrients*, Macroalgae,** Dissolved oxygen
	Sediments, Metals	**TSS, Water nutrients***	Dissolved oxygen
Good sediment (seabed) quality	Nutrients, Metals	**Sediment metals, nutrients**	TOC, RDP, sulphides
	Sediments	Extent of dominant substrate types	**Mud content, Deposition rate**
Diverse and high-quality habitats	Nutrients, Sediments	Extent of habitats	**Macroalgae, macrofauna,** shellfish
Healthy levels of primary production (non-eutrophic)	Nutrients, Sediments	**Macroalgae, sediment chl-a**	
Diverse and functional faunal communities	Nutrients, sediments, metals	**Macrofauna**	**Shellfish metals**
Healthy fish populations	Harvesting, Nutrients, sediments, metals	Finfish diversity and abundance	
Human health			
Uncontaminated water	Faecal bacteria, metals	**Water FIB**	Frequency of bathing beach closures
Uncontaminated shellfish	Faecal bacteria, metals	**Shellfish FIB and metals**	Frequency of harvest closures
Mahinga kai			
Harvest accessibility	Sediments, Metals, faecal bacteria	**TSS, mud content, Water FIB**	
Plenty to harvest	Nutrients, Sediments	**Macrofauna** and finfish abundance	
Safe to eat	Metals, faecal bacteria	**Shellfish FIB and metals**	

Only three variables not previously prioritised as attributes were identified as the most direct measures of a stressor: "finfish diversity and abundance", "extent of dominant substrate types" and "extent of habitats". Finfish was not included in the prioritised list of attributes, as the predominant stressor (commercial and recreational harvesting) does not originate upstream. The dominant substrate type that is primarily affected by upstream management is extent of muddy substrates and this was included as a prioritised attribute. Extent of habitats was considered to be well covered by macroalgae and macrofauna. Only seagrass was not covered - many experts considered the differences in temporal dynamics exhibited by seagrass cover between north and south regions sufficient to preclude inclusion of seagrass as a national attribute or state variable.

The only proposed attributes that did not represent a direct measure of values were water chl-a and deposition rate.

4.3 Robustness and consistency of methods

The degree to which robust, standardised methods are used for both collection and analysis of data was considered to be of secondary importance. Accordingly, we only discuss those proposed variables considered to be predictable from upstream measures and with direct links to values in this stage. We removed variables that experts agreed were currently not well predicted by upstream measures (rank 5 or 6, water and sediment chl-a and shellfish metals), as well as those that did not represent a direct measure of a value (water chl-a and deposition rate). These, however, are discussed further in section 4.5.1.

Initially, we considered only the information that had been collected as part of activities described in Section 3.1 and Section 3.2. However, following further review of acquired information, we realised that some of the problems identified in the existing data could be solved relatively simply in Stage 2 (Table 4-3). Other problems would however require alteration of methods and further research (leading to solutions over a longer-term); in some instances, no clear solution was foreseen.

Table 4-3: Potential bottlenecks identified and suggested solutions (either in Stage 2 or long-term).

Attributes	Potential bottlenecks	Easy fixes in Stage 2[15]	Longer term fixes	New research needed
Water nutrients	Water column nutrient concentrations do not necessarily reflect the quantity of nutrients available to primary producers	Analyse concentrations vs loadings		
	Variable analytical methods with limited comparability		Standardise (see NEMS recommendations)	
Water clarity	Variable analytical methods with limited comparability		Standardise (see NEMS recommendations)	
TSS	Spot sampling for TSS is likely to be negatively impacted by the high temporal variation in suspended sediment loads		Assess cost-effectiveness of continuous sampling	
	Variable analytical methods with limited comparability		Standardise (see NEMS recommendations)	
Water FIB	Inappropriate sampling design, in terms of spatial extent and number of replicates, non-representative sampling sites	Analyse effects of replication and extent	Standardise nationwide, ultimately – NEMS development	
	Inconsistency in metrics and statistics used		Standardise nationwide, ultimately – NEMS development	
	Contamination of samples during lab incubation resulting in false alarms	Provide recommendations on methods used	Quality assurance	
Shellfish FIB	Inappropriate sampling design, in terms of spatial extent and number of replicates, non-representative sampling sites	Analyse effects of replication and extent	Assess whether guidelines that meet export requirements are appropriate for recreational users	
	Samples are often biased toward good weather conditions and when shellfish are being harvested			Yes
	Potential false alarms due contamination		Quality assurance	

[15] Easy fixes refer here to the already available datasets – what can be done to improve their usability in Stage 2, while longer term fixes are suggested for future considerations to improve monitoring of attributes and state variables and ensure quality of acquired information.

Attributes	Potential bottlenecks	Easy fixes in Stage 2[15]	Longer term fixes	New research needed
	issues during incubation			
	Time-consuming lab analyses (culture-based) do not allow "real time" response			Develop faster FIB assessment methods, molecular techniques
Macroalgae	Lack of adequate training for consistent assessment of macroalgae		Quality assurance	
	Uncertainty or inconsistency in the ground-truthing undertaken in different estuaries		Standardise nationwide, ultimately – NEMS development	
Macrofauna	Selection non-representative sampling sites	Analyse to determine best sites	May need new sites	
	Limited, variable and inconsistent taxonomic resolution	Standardise, e.g., by using coarser taxonomic resolution when taxa inconsistently identified	Quality assurance and national taxonomic database; use of molecular ID methods	Development of standardised molecular-based ID guidelines
Sediment Chl-a	Inappropriate sampling design, in terms of spatial extent and number of replicates, non-representative sampling sites	Analyse effects of replication and extent		
	Inconsistent analytical methods, producing often incomparable results		Standardise nationwide, ultimately – NEMS development	
	Inconsistent units used in reports	Conversion to standard		
Mud content/ grain size	Inappropriate sampling design, in terms of spatial extent and number of replicates, non-representative sampling sites	Analyse effects of replication and extent	Standardise nationwide, ultimately – NEMS development	
	Differences in sample depth		Standardise nationwide, ultimately – NEMS development	
	Results between and within two main grain size analysis methods are not necessarily comparable	Analyse for effect of difference	Standardise nationwide, ultimately – NEMS development	

Attributes	Potential bottlenecks	Easy fixes in Stage 2[15]	Longer term fixes	New research needed
	Inconsistent size fractions are used for reporting mud often make results incomparable		Standardise nationwide, ultimately – NEMS development	
Sediment deposition rate	Inappropriate sampling design, in terms of spatial extent and number of replicates, non-representative sampling sites	Analyse effects of replication and extent	Standardise nationwide, ultimately – NEMS development	
	For modelled information, poor consensus on which national scale models to use to estimate sediment loads		Standardise nationwide, ultimately – NEMS development	
	Three main methods are presently used	Assess comparability of results obtained by different methods	Standardise nationwide, ultimately – NEMS development	Development of a standardised technique
	Metrics unrelated to specific estuary conditions		Use of multiple complementary methods (fine-scale and broad-scale) to increase confidence and account for spatial variability	
	Inappropriate sampling design, in terms of spatial extent and number of replicates, non-representative sampling sites	Analyse effects of replication and extent		
Sediment metals	Differences in sample depth		Standardise nationwide, ultimately – NEMS development	
	Inconsistent analytical methods	Assess comparability of results obtained by different methods	Standardise nationwide, ultimately – NEMS development	
	Inappropriate sampling design, in terms of spatial extent and number of replicates, non-representative sampling sites	Analyse effects of replication and extent	Standardise nationwide, ultimately – NEMS development	
Shellfish metals and other contaminants	Inappropriate sampling design, in terms of spatial extent and number of replicates, non-representative sampling sites	Analyse effects of replication and extent	Standardise nationwide, ultimately – NEMS development	
	Time-consuming lab analyses (culture-			Develop faster FIB assessment methods,

Attributes	Potential bottlenecks	Easy fixes in Stage 2[15]	Longer term fixes	New research needed
	based) do not allow "real time" response			molecular techniques
	Limited number of contaminants with human health standards			Human health-contaminant guidelines
	Some chemical contaminants of potential concern may have limited bioaccumulation - difficult to detect			Bioaccumulation studies
	Varying detection limits between laboratories		Quality assurance	

For ongoing analysis, the responses from technical experts regarding the possibility of using a single robust method were used. We had queried the use of a single method for both candidate attributes and state variables. Responses were ranked from 1 (experts agreed there was a robust method) through to 6 (experts agreed that there was no robust method), with ranks of 3 and 4 indicating lack of consensus amongst experts (Table 4-4). To keep this section useful for future reference, we assessed all proposed attributes and state variables from the Stage 1A report.

Experts agreed that robust methods generally existed for the following attributes and state variables:

Proposed attributes (rank 1 and 2):	Proposed state variables (rank 1 and 2):
Total suspended solids	Dissolved oxygen
Water Chl-a	Frequency of bathing beach closures
Water faecal indicator bacteria	Frequency of harvest closures
Shellfish faecal indicator bacteria	Sediment nutrients (TN, TP)
Macrofauna	Sediment sulphides
Sediment metals	Extent of habitats
Shellfish metals	

Table 4-4: Variables ranked by robustness of method. 1 = agreement that method is robust, 3 – 4 = disagreement about whether the method is robust, 6 = agreement that a robust method does not exist.

Attributes/ State variables		Rank assigned
Attributes	(d) Total suspended solids	1
	(b) Water Chl-a	2
	(e) Water faecal indicator bacteria	2
	(m) Shellfish faecal indicator bacteria	2
	(g) Macrofauna	2
	(k) Sediment metals	2
	(l) Shellfish metals	2
	(a) Water nutrients (TN, TP)	3
	(c) Water clarity	3
	(h) Sediment Chl-a	3
	(i) Mud content/grain size	3
	(f) Macroalgae	4
	(j) Sediment deposition rate (incl. AASR)	6
State variables	(n) Dissolved oxygen	1
	(w) Frequency of bathing beach closures	1
	(x) Frequency of harvest closures	1
	(o) Sediment nutrients (TN, TP)	2
	(q) Sediment sulphides	2
	(t) Extent of habitats (e.g., seagrass beds)	2
	(p) Sediment TOC	3
	(s) Extent of dominant substrate types (e.g., mud)	3
	(y) Harvest area accessibility	3
	(v) Shellfish distribution and abundance	4
	(r) Redox potential discontinuity depth	5
	(u) Finfish diversity and abundance	6

4.4 Understood or predictable variability

The final important aspect considered was whether large-scale temporal and spatial variability was generally understood. This allows us to determine whether such variability will limit the usefulness of a variable as either an attribute or as a state variable. If temporal variability was adequately understood it would be possible to derive bands that incorporate climatic variability; if spatial variability were adequately understood it would allow us to create bands separately for different estuary types, west coast vs east coast or north vs south, as required. Rankings similar to those used in the previous sections were assigned, ranging from rank 1 (experts agreed variability was understood and predictable) through to 6 (experts agreed that variability was not understood), with ranks of 3 and 4 showing lack of consensus amongst experts (Table 4-5).

Experts generally agreed that temporal variability would not preclude development of the variables as attributes for:

- Sediment deposition rate.
- Water nutrients (TN, TP).
- Macroalgae.
- Macrofauna.
- Mud content/grain size.
- Sediment metals.

However, definite problems were identified for "water faecal indicator bacteria" and "shellfish metals", driven by poor understanding of what was likely to cause temporal variation. For state variables, no problems were identified related to temporal variability for the following variables:

- Extent of dominant substrate types.
- Frequency of bathing beach closures.
- Harvest area accessibility.
- Dissolved oxygen.
- Frequency of harvest closures.

Limitations were identified for "sediment sulphides", and "finfish diversity and abundance".

Experts agreed that more was known about large-scale temporal variability than large-scale spatial variability (i.e., differences driven by estuary type, between coasts and longitudinal gradients, see Table 4-5). We anticipated that much of this information would be analysed in Stage 2, so we assessed the likelihood that the information required could be provided in Stage 2 using information regarding data availability for each variable. This was coded as "fixable in Stage 2 analysis", "yes", "probable" or "requires new data collection" (Table 4-5).

Information regarding the potential time lags between changes made in upstream catchments (e.g., leading to reduced TN concentrations in streams entering the estuary), and responses observed in the estuary could also be considered (Table E-1, Appendix E). The confidence intervals around these are large and this uncertainty has led us to not consider them further when making the recommendations.

Table 4-5: Consensus on the spatial and temporal variability of proposed variables.

	Attribute/ State variables	Temporal variability rank	Spatial variability rank	Fixable in Stage 2 analysis	Requires further data collection
Attributes	Sediment deposition rate (incl. AASR)	1	4		yes
	Water nutrients (TN, TP)	2	5	yes	
	Macroalgae	2	6		yes
	Macrofauna	2	4	probably	
	Mud content/grain size	2	3	probably	
	Sediment metals	2	3	probably	
	Water Chl-a	3	6	probably	
	Water clarity	3	5	probably	
	Sediment Chl-a	3	6		yes
	Total suspended solids	4	6	probably	
	Shellfish faecal indicator bacteria	4	5	probably	
	Water faecal indicator bacteria	5	6	yes	
	Shellfish metals	5	4		yes
State variables	Extent of dominant substrate types	1	4		yes
	Frequency of bathing beach closures	1	2		
	Harvest area accessibility	1	1		
	Dissolved oxygen	2	5	yes	
	Frequency of harvest closures	2	2		
	Sediment nutrients (TN, TP)	3	5	probably	
	Sediment TOC	3	5	probably	
	Redox potential discontinuity depth	3	4	yes	
	Extent of habitats	4	5		yes
	Shellfish distribution and abundance	4	3		yes
	Sediment sulphides	5	5		yes
	Finfish diversity and abundance	6	6	probably	

4.5 Summary and recommendations

4.5.1 State variables

Results from Section 3.3 highlight that all proposed attribute variables are also suitable for consideration as state variables (SVs). However, although all these variables may serve as SVs, their importance in terms of linkage to values differs. They also differ in terms of the robustness of currently used sampling techiques, currently available analytical methods, and the degree to which we understand their natural variability. To prioritise SVs for development in Stage 2, we recommend that their importance in terms of linkages to values, the amount of work required for method development, the need for further data collection to determine status and baselines, practicality (requires input from RCs), and required alignment with other variables for interpreting results, are considered further. Table 4-6 summarises these factors for all attributes considered for inclusion as

SVs, providing a starting point for these considerations. Finfish abundance and diversity is omitted from this consideration - although it is strongly linked to both ecosystem health and mahinga kai, data are rarely collected owing to the high survey cost. No data less than 10 years old are available, so, finfish abundance and diversity could not form part of a Stage 2 evaluation of estuarine status.

When considering standardisation of methods it is important to determine whether the inconsistencies are minor (e.g., slight differences in the size of sampler), and whether standardisation is required (e.g., using weak or strong acid extraction). When the inconsistency is due to differences in where and when past data collection has been carried out during different studies, standardisation would need to be carefully considered against the sampling question and the value of any existing time series. For example, should standardised samples for present status be taken: randomly; in a representative habitat; where physical or chemical change is most likely to occur; or where the ecological response is most likely to be observed? This would require adequately understanding:

- the rationale for selection of existing locations,
- the purpose of the study, and
- the effect standardisation would have on ongoing studies.

Referring to the information in Table 4-6, the level of effort to develop SVs ranges from those requiring major work and extensive data collection (1), to those requiring least work (9 - 11) as follows:

1. New data and research required (Shellfish metals);
2. Links to ecosystem health to be established, new data and models or method development (Sediment deposition). *Note*: *this is not a direct measure of ecosystem health*;
3. New data and links to ecosystem health to be established and a change in method may be required (Mud content). *Note*: *this is not a direct measure of ecosystem health*;
4. Links to ecosystem health to be established and this would have to be done in conjunction with other SVs (Sediment nutrients (TN, TP), Sediment TOC, Water nutrients (TN, TP), Water Chl-a concentration, Water clarity, Total suspended sediments). *Note: of these variables, water clarity is the most direct measure of ecosystem health and Water Chl-a and sediment TOC are the least direct measures;*
5. Development and/or confirmation of standard metrics (Water faecal indicator bacteria, Shellfish faecal indicator bacteria). *Note: these are both direct measures of human health and mahinga kai*;
6. New data and requires standard metrics for water and shellfish faecal indicator bacteria to be developed (Frequency of bathing beach closures, Frequency of harvest closures, Harvest area accessibility);
7. New data and links to ecosystem health or mahinga kai to be determined (Sediment Chl-a concentration, Extent of habitats, Extent of dominant substrate types, Shellfish distribution and abundance). *Note: of these variables, Extent of habitats, Extent of dominant substrate types and Shellfish distribution and abundance are the most direct measures of ecosystem health;*

8. New data to be collected to validate overseas guidelines, but also need other SV to be fully interpretable (Sediment sulphides, Redox potential discontinuity depth). *Note: these are not direct measures of ecosystem health*;

9. New data and validation of guidelines (Macroalgae, Sediment metals). *Note: these are both direct measures of ecosystem health*;

10. Minor analyses to determine what degree of standardisation is necessary for sampling design and comparison of present metrics (Macrofauna). *Note: this is a direct measure of ecosystem health and mahinga kai*;

11. Minor analyses but requires measurement of other variables to be fully interpretable (Dissolved Oxygen). *Note: this is not a direct measure of ecosystem health (except at its extremes)*.

Table 4-6: Relationship between variables with potential to serve as state variables (SVs) and factors likely to determine their future usefulness as SVs.

Variable	Importance[16] / *link to values*	Issues to be resolved to improve robustness[17]	Easy fixes in Stage 2	Longer term fixes	Alignment with other variables
Water nutrients (TN, TP)	Medium / Clean uncontaminated water, *Ecosystem health*	Inconsistency in analytical methods and sampling design	Determination of "best" method and recommendations for standardisation	Assess cost-effectiveness of continuous sampling	Should be complemented by other water quality variables for correct interpretation
Water Chl-a concentration	Low / Clear water, *Ecosystem health*	Inconsistency in analytical methods and sampling design	Recommendations for standardisation	Assess cost-effectiveness of continuous sampling	Should be complemented by water nutrients and clarity for correct interpretation
Water clarity	High / Clear water, *Ecosystem health*	Inconsistency in analytical methods and sampling design	Determination of "best" method and recommendations for standardisation	Assess cost-effectiveness of continuous sampling	Should be complemented by Total suspended sediment information for correct interpretation
Total suspended sediments*[18]	Medium / Clear water, *Ecosystem health*	Inconsistency in analytical methods and sampling design	Determination of "best" method and recommendations for standardisation	Assess cost-effectiveness of continuous sampling	
Dissolved oxygen	Low / Clean uncontaminated water, *Ecosystem health*	Inconsistency of sampling design	Recommendations for standardisation	Assess cost-effectiveness of continuous sampling	For correct interpretation, should be measured alongside other variables (TSS, water nutrients, physical parameters, Chl-a, sediment RDP, etc.)
Macrofauna*	High / Diversity of fauna, *Ecosystem health*	Inconsistency of sample location,	Assessment of effect of inconsistencies,	Possible need to develop rationale for selection of new sites	

[16] Importance was assessed as "high" if a variable represents the only direct measure of major value/objective-related stressors, "medium" – if a variable identified as one of the direct measures of major value/objective-related stressors, and "low" – if a variable is listed among other relevant measures (see Table 4-2).
[17] More details on monitoring methods and inconsistencies are given in Appendix B.
[18] Those marked with an asterisk are variables recommended for further development as attributes in Stage 2; they are considered here as state variables for monitoring the state of estuary values (see Table 4-3 for information on attributes' potential bottlenecks and suggested solutions)

Variable	Importance[16] / *link to values*	Issues to be resolved to improve robustness[17]	Easy fixes in Stage 2	Longer term fixes	Alignment with other variables
		taxonomy and analysis	recommendations for standardisation		
Macroalgae*	Medium / Diversity of flora, *Ecosystem health*	Inconsistency in ground truthing and assessment	Recommendations for standardisation	Need to determine practical lowest resolution before collecting new data from northern areas	
Sediment deposition rate (incl. AASR)*	Low / Diversity of substrate types, *Ecosystem health*	Inconsistency in sampling, little data, often modelled		Need separate model for each estuary, need to train models with 'event' based data or develop new method	
Mud content/ grain size*	Low / Diversity of substrate types, *Ecosystem health*	Inconsistency of sampling design, three methods used	Analyse for effects of inconsistencies in sample design	Analyse differences between methods, recommendations for standardisation likely to require additional data collection and analysis	
Sediment metals*	Medium / Uncontaminated sediment, *Ecosystem health*	Inconsistency of sampling design and analyses, Low to moderate data available	Analyse for effects of inconsistencies in sample design, recommendations for standardisation	New data would need to be collected around NZ to validate present guidelines	
Sediment Chl-a concentration	Medium / Diversity of flora, *Ecosystem health*	Inconsistency of sampling design and analyses	Analyse for effects of inconsistencies in sample design, recommendations for standardisation	Need for new data and analysis to establish links to ecosystem health status	
Sediment nutrients (TN, TP)	Medium / Uncontaminated	Inconsistency of laboratory analyses	Recommendations for standardisation	Establish links to ecosystem health status	Should be complemented by other sediment quality variables for correct interpretation

Variable	Importance[16] / *link to values*	Issues to be resolved to improve robustness[17]	Easy fixes in Stage 2	Longer term fixes	Alignment with other variables
	sediments, *Ecosystem health*				
Sediment TOC	Low / Uncontaminated sediments, *Ecosystem health*	Inconsistency of sampling design and non-standardised use of surrogate measures	Recommendation for standardisation	Establish links to ecosystem health status- likely to be difficult as has not been achieved overseas	Should be complemented by other sediment quality variables for correct interpretation
Sediment sulphides	Low / Uncontaminated sediments, *Ecosystem health*	Inconsistent sampling and analytical approaches	Recommendations for standardisation and thresholds	New data are needed to validate overseas guidelines for different estuary typologies and geology	Should be complemented by other sediment quality variables for correct interpretation and biological information
Redox potential discontinuity depth	Low / Uncontaminated sediments, *Ecosystem health*	Inconsistency in sampling design, Low to moderate data only available	Recommendations for standardisation	New data are needed to validate overseas guidelines for different estuary typologies and geology	Should be complemented by other sediment quality variables, geochemistry and biological information for correct interpretation
Extent of habitats	High / Diversity of habitats, *Ecosystem health*	Little to moderate data only available, inconsistent sampling approaches	Recommendations for standardisation	New data are needed to establish links to ecosystem health status	
Extent of dominant substrate types	High / Diversity of substrate types, *Ecosystem health*	Little to moderate data only available, inconsistent sampling approaches	Recommendations for standardisation, cost-benefit analysis	New data are needed to establish links to ecosystem health status	
Water faecal indicator bacteria*	High / *Human health*	Inconsistencies in methods, sampling design, metrics and statistics	Analyse effects of inconsistencies, Recommendations for standardisation of sampling	Recommendations for standard metrics	

Variable	Importance[16] / *link to values*	Issues to be resolved to improve robustness[17]	Easy fixes in Stage 2	Longer term fixes	Alignment with other variables
Shellfish faecal indicator bacteria	Medium / *Human health*	Inconsistencies in sampling design, metrics and statistics	Analyse effects of inconsistencies, Recommendations for standardisation of sampling	Recommendations for standard metrics, Assess whether export requirements are appropriate for recreational users	
Shellfish metals	Medium / Uncontaminated sediment, *Ecosystem health, Human health*	Inconsistencies in sample design, varying bioaccumulation rates between species and chemicals, varying detection limits	Analyse effects of inconsistencies, recommendations for standardisation and quality assurance	Need for new data, studies on bioaccumulation rates and development of health guidelines	
Shellfish distribution and abundance	High / Diversity of habitats, *Ecosystem health, Mahinga kai*			New data are needed to establish links to ecosystem health status	
Frequency of bathing beach closures	Low / *Mahinga kai*	Little data available		New data are needed to establish mahinga kai guidelines	
Frequency of harvest closures	Low / *Mahinga kai*	Little data available		New data are needed to establish mahinga kai guidelines	
Harvest area accessibility	Medium / *Mahinga kai*	Little data available		New data are needed to establish mahinga kai guidelines	

*Variables recommended for further development as attributes in Stage 2, considered here for serving as state variables (see Table 4-3 for full information on attributes' potential bottlenecks and suggested solutions)

4.5.2 Attributes

The results of the rankings and considerations for attributes are summarised in Figure 4-1, providing a shorter list of overall best-performing candidates (selected in at least 3 of the 4 filters applied).

The results of the rankings of the attributes proposed in the Stage 1A report were discussed at a meeting with MfE and their advisory panel (7th December 2017), with the following outcomes:

- Water column nutrients and sediment deposition rate were both separated into two variables (measured and modelled variables) and re-assessed for their ability to meet the four important aspects of an attribute (predictability from upstream measures, linkage to a value, robust methods and low or predictable temporal variability.
- Predictability from upstream measures were re-assessed by suitably experienced experts present at the meeting to better reflect the degree of uncertainty generally accepted by the freshwater National Objectives Framework. Experts were:
  - Water column (Chris Cornelisen).
  - Macrofauna (Judi Hewitt).
  - Macroalgae (Judi Hewitt, Ton Snelder).
  - Sediment characteristics and shellfish metals (Megan Carbines).
  - FIB (Rebecca Stott).
  - Suspended sediment and sediment deposition rates (Mal Green).
- Ranking of robust methods was discussed and re-considered by the technical experts to better reflect our ability to standardise methods with currently available data.

As a result of this discussion, three variables were identified as strongly predictable by upstream measures:

- Modelled water nutrient concentrations (TN, TP).
- Modelled AASR.
- Measured sediment deposition rate.

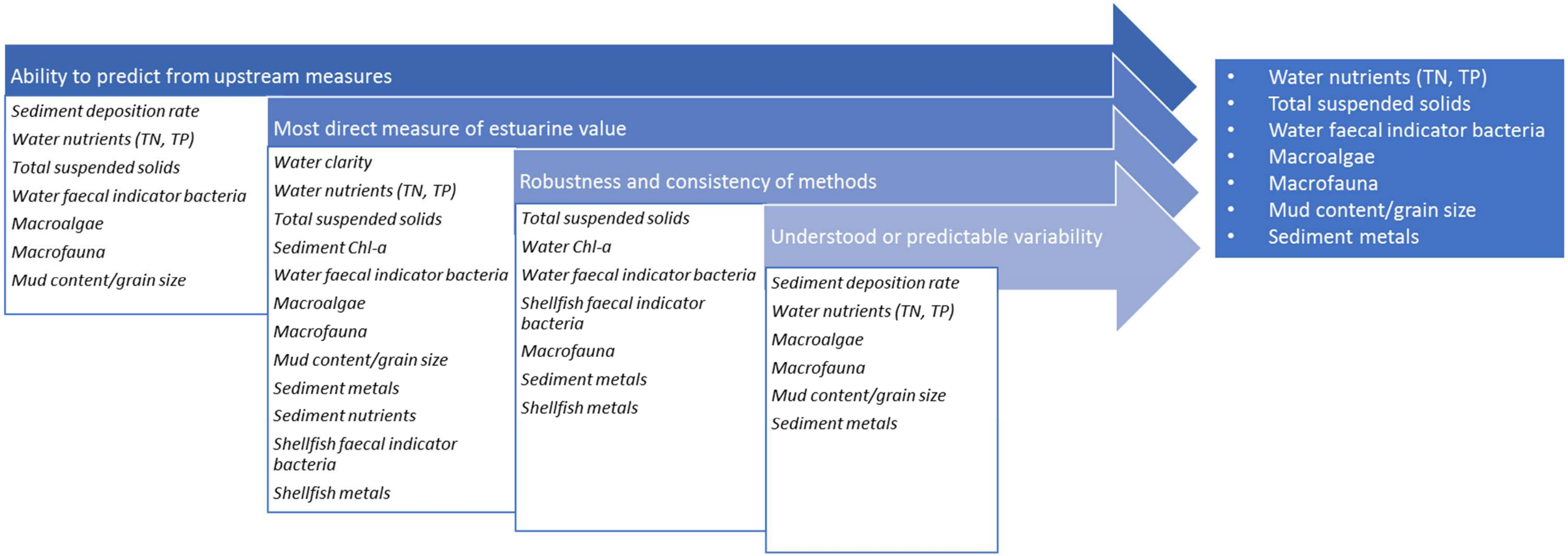


Figure 4-1: Overall results of attribute filtering based on the outcomes of this stage of the project. Attributes rated highly in at least 3 categories were short-listed for further consideration.

Unfortunately, the link between modelled water nutrient concentrations and ecosystem health values was unknown, and deposition rate (whether predicted or measured) was not considered to have a strong direct link to ecosystem value.

The three variables considered next most likely to be predicted by upstream measures were:

- Macroalgae (these have been strongly linked to TN and TP loads);
- Mud content (predicted using catchment and hydrodynamic models);
- Macrofauna (very recent analysis in two estuaries has demonstrated that it is possible to predict different species abundances from sediment concentrations or yields and TN concentrations).

These three variables are all direct measures of ecosystem health or mahinga kai. Another four variables were suggested to have the potential to be predictable from upstream measures:

- Total suspended sediments.
- Water FIB.
- Sediment metals.
- Measured water nutrients (TN, TP).

Any of the seven variables considered to be reasonably predictable using upstream measures, and that are directly related to estuarine values, would make good attributes. We then considered the three major stressors (sediments, nutrients and faecal contaminants), the three values (ecosystem health, human health and mahinga kai), and the state of previous work on limits and guidelines for estuaries. The following conclusions were drawn:

1. The sediment guidelines for estuaries state clearly that it would be important to consider deposition rate, extent of mud content and suspended sediment concentrations if estuaries were to be protected from sediment inputs; this suggests that work that leading to the development of a combined objective and thresholds would be useful.
2. The ETI has created a strong endpoint relating the effects of nutrients on Macroalgae. This is very close to being a fully developed attribute.
3. Sediment metals are predictable and link well to ecosystem health, so they could form a good attribute; it may also be possible relate these to mahinga kai, although few specific and relevant human health guidelines exist.
4. Of the two attributes proposed for human health, Water FIB is preferred. Ongoing research in several projects, as well as analysis that will be conducted in the monitoring status section of Stage 2 could improve our ability to develop a strong attribute.
5. Macrofauna are a primary indicator of estuarine ecosystem health and are likely to be particularly useful for describing a series of states between pristine and poor health. Different macrofaunal species respond uniquely to sediments, nutrients and heavy metals, mainly because of their biological traits – this creates the potential that we may develop a set of attributes able to discriminate the influence of these three stressors across New Zealand, independent of estuary typology.

At this stage, we are not recommending the development of an attribute around water column nutrients. Development of such an indicator may however be necessary to create objectives for responses to nutrients before excessive macroalgal growth occurs. Another five of the proposed attributes were not considered at this time to be predictable enough by upstream measures for further use:

- Shellfish faecal indicator bacteria.
- Shellfish metals.
- Sediment Chl-a.
- Water clarity.
- Water Chl-a.

Most of these are, however, direct measures of ecosystem health, human health or mahinga kai, and it may be necessary to develop them as attributes in future to fully protect estuaries.

Finally, we recommend that process variables (e.g., denitrification), emerging contaminants (e.g., plastics) and molecular FIB markers are considered in the future - improved understanding of their importance, as well as improvement and more widespread availability of monitoring techniques is expected.

The linkages between attributes selected for development in Stage 2 of the project, estuary values, and aspects to be managed are presented in Figure 4-2.

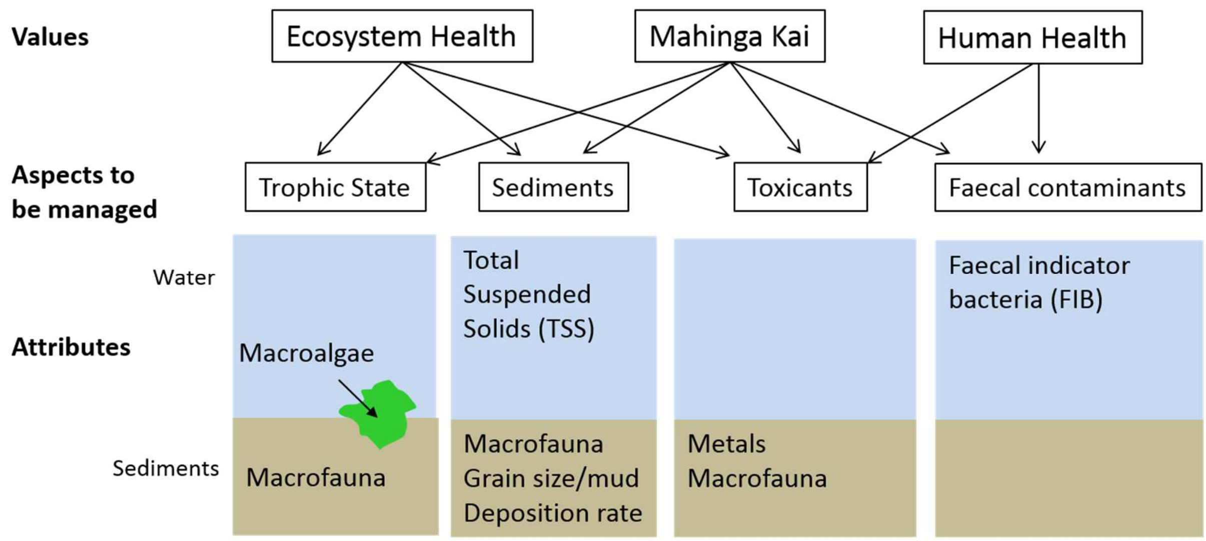


Figure 4-2: Variables recommended for further development as attributes in Stage 2 of the project.

5 Acknowledgements

We thank the many people from the wider project team and beyond who participated in the on-line expert survey and follow up consultation and the evaluation of variables (see Appendix C). We also wish to thank internal reviewers of earlier drafts, including Anna Madarasz-Smith (Hawke's Bay Regional Council), Megan Carbines (Auckland Council) and Helli Ward (MfE). Constructive external reviews provided by Ton Snelder and Peter Scanes were greatly appreciated. Input by these individuals as well as Mal Green (Streamlined Environmental) and Helen Kettles (Department of Conservation) during the MfE review meeting was very helpful in focusing efforts and finalising this report.

MfE thanks members of the External Advisory Group: Helen Kettles (Department of Conservation), Megan Oliver (Greater Wellington Regional Council), Marcus Cameron (Auckland Council), Rob Smith (Tasman District Council) and Rich Ford (Ministry for Primary Industries) for their participation.

This project follows on from the formative work on an Estuary National Objectives Framework initiated in 2012. We acknowledge Helen Kettles and Paula Warren (Department of Conservation) for leading this work, with support from John Phillips (Ministry for the Environment). MfE also thanks the experts who took part in that early process: John Zeldis, Malcolm Green and Drew Lohrer (NIWA); Joanne Ellis (Cawthron Institute); Leigh Stevens (Wriggle Coastal Management); and Anna Madarasz-Smith (Hawke's Bay Regional Council).

6 Glossary of abbreviations and terms

The following table provides definitions and narratives for a range of terms used in this project. For consistency, we have incorporated wording and definitions from the NPS-FM.

Abbreviation/ term	Explanation
AASR	Annual average sedimentation rate.
Areal extent	The extent of a 2-dimensional surface enclosed within a specified boundary. Measures of areal extent of habitats are typically determined with the aid of aerial imagery and walking the estuary to delineate areas with images, maps and a GPS.
Attribute	Measurable characteristics of estuaries, including physical, chemical and/or biological, properties that are directly affected by upstream aspects to be managed, such as sediments and nutrients.
Biogenic habitat	Biogenic habitats are created by plants and animals and may be the organism itself, such as a seagrass meadow or a bed of horse mussels, or arise from an organism's activities, such as the burrows created by crabs. Examples in New Zealand estuaries include mangrove forests, seagrass meadows, green-lipped mussel and oyster reefs. Less widely recognised examples are horse mussel beds, bryozoan fields, tubeworm mounds, dog cockle beds, and beds of *Caulerpa*, a green alga.
Chl-a	Chlorophyll-a.
CLUES	Catchment Land Use for Environmental Sustainability is a GIS-based modelling system which assesses the effects of land use change on water quality and socio-economic indicators.
Coastal hydrosystem	A coastal system comprising hydrological, geomorphic and ecological components, including significant surface water and/or groundwater components, that spans within a gradient through fresh water to brackish to saline (Hume et al. 2016).
Coastal marine area	The foreshore, seabed, and coastal water, and the air space above the water: a) of which the seaward boundary is the outer limits of the territorial sea: b) of which the landward boundary is the line of mean high water springs, except that where that line crosses a river, the landward boundary at that point shall be whichever is the lesser of— (i) 1 kilometre upstream from the mouth of the river, or (ii) the point upstream that is calculated by multiplying the width of the river mouth by 5. (RMA definition).
Coastal water	Means seawater within the outer limits of the territorial sea and includes: a) seawater with a substantial freshwater component, and b) seawater in estuaries, fiords, inlets, harbours, or embayments. (RMA definition).
Community	An assemblage of two or more species of organisms and/or populations interacting in a specific area (habitat) or time.
EC	Emerging contaminant(s).
EMP	Estuary Monitoring Protocol – a national protocol prepared for estuarine environmental assessment and monitoring to support councils and the Ministry for the Environment.
Estuary	Estuaries are spatially bounded as seaward from an imaginary line closing the mouth (opening to the ocean), to landward where ocean derived salts measure less than 0.5ppt during the period of average annual low flow (Robertson et al. 2016a). The

Abbreviation/ term	Explanation
	recent coastal hydrosystems typology defines an estuary as partly enclosed by land, open to the sea for extended periods, within which seawater is measurably diluted by land drainage, and which typically experiences daily tidal ingress (i.e., has a tidal prism; Hume et al. 2016).
Eutrophication	Process whereby excessive nutrient inputs to a water body result in accelerated primary production (phytoplankton and macroalgae growth), and flow-on effects to the wider ecosystem, such as reduced water clarity, physical smothering of biota, or extreme reductions in dissolved oxygen because of microbial decay.
FIB	Faecal indicator bacteria - types of bacteria used to detect and estimate the level of faecal contamination of water.
Flushing	Using measures of tidal range, and the ratio of river runoff to estuarine volume, flushing is the time for freshwater inflows and the tidal prism volume to replace the estuary volume. An estuary with large volumes and short flushing times are less susceptible to eutrophication from upstream nutrient loading than estuaries with smaller volumes and long flushing times.
Habitat	An ecological area made up of physical and biological factors that provides an organism(s) with food, shelter, ability to reproduce, etc.
Inorganic compounds	Any compound that lacks a carbon atom and is not of biological origin. For example, trace metals, minerals and inorganic forms of nutrients.
LAWA	Land Air Water Aotearoa.
Limit	Based on the NPS-FM definition, a limit is the maximum amount of resource that is available for use while still enabling an objective to be met. It is a specific quantifiable amount that links the objective (the desired state) to use of the resource. A limit puts constraints on how much of that resource is available for use. As an example, for estuary *water quality*, the assimilative capacity of the water (its ability to absorb contaminants) is the resource being limited. A quality limit would describe how much of a contaminant (e.g., a nutrient) could be discharged into the water by users without exceeding an objective.
Macrofauna	Macrofauna are invertebrates that live on or in sediment, or attached to hard substrates. They include infauna (those in the sediments) and epifauna (those colonising the surface of sediments). They are generally classified according to size, with invertebrates greater than 0.5 mm or 1 mm in size regarded as macrofanua.
Mahinga kai	Māori traditional food species gathered from the environment. The definition also includes the places these species are gathered and the practices involved in their collection. Indigenous estuarine species have traditionally been used as food, tools, or other resources.
National bottom line	Based on the NPS-FM definition, the national bottom line is the boundary between the C and D states for the attributes associated with the compulsory national values (‘ecosystem health’ and ‘human health for recreation’). According to this definition, all estuaries (or manageable units within estuaries) would have objectives set above nationally-defined bottom lines.
National Objectives Framework (NOF)	The National Objective Framework (NOF) directs regional decision-making in the setting of objectives. It consists of a process, a set of national values, and a set of attributes for setting freshwater objectives to achieve those values.
National value	Originating from the NPS-FM, national values are those intrinsic qualities, uses or potential uses that were determined by Government both to be appropriate based on

Abbreviation/ term	Explanation
	a set of criteria, and to be of national significance. Some are compulsory and must have objectives set for them, while others may be considered compulsory at a regional level by regional councils.
Naturally occurring processes	Processes that could have occurred in New Zealand prior to the arrival of humans. In the case of the NPS-FM, where existing conditions are below a national bottom line due to naturally occurring processes, a regional council may set an objective below a national bottom line. By definition, any deterioration in water quality that is caused by human interventions, and would not have occurred without that intervention, does not qualify a water body to have an objective set for it below a bottom line.
NEMS	National Environmental Monitoring Standards.
NPS-FM	National Policy Statement for Freshwater Management.
Organic compounds	Organic compounds contain carbon atoms and can be of synthetic or natural origin. Those that can be toxic to organisms include compounds derived from petroleum and gas (polycyclic aromatic hydrocarbons, or PAHs), organic herbicides, and organochlorine insecticides.
OTOT	Oranga Taiao, Oranga Tangāta - Knowledge and toolsets to support co-management of estuaries, a four-year MBIE research programme.
Pressure	Pressures are the human activities (e.g., urbanisation, farming, climate change) and natural processes (e.g., floods) that generate stressors that in turn lead to environmental changes.
REDOX	Reduction-oxidation potential, a measure of the reducing conditions in a medium, e.g., sediment.
RPD - REDOX Potential Discontinuity	The zone within estuarine sediments where it changes from aerobic to anaerobic conditions. It can be visually assessed by observing the colouration gradient of well oxygenated sediments near the surface (lightly coloured) to anaerobic sediments (black) that are deeper within a collected core sample.
Secondary contact	People's contact with water that involves only occasional immersion and includes wading or boating (except boating where there is high likelihood of immersion; NPS-FM definition). The term is used in relation to objectives that require the health of people and communities, at least as affected by secondary contact with water, to be safeguarded. This objective is supported by the compulsory national value 'human health for recreation'.
SS	Suspended sediment.
State Variable	Measurable variables or metrics derived from multiple variables that provide information about and/or describe the state of estuary values.
Stressor	Stressors are the physical, chemical, or biological 'agents of change' on ecosystem health, functioning and productivity or human health for recreation, or mahinga kai. Sediment loading is an example of an upstream stressor that affects estuaries.
Substrate	The sediment or material on or from which an organism grow and live.
Taonga species	Species of native birds, plants and animals of special cultural significance and importance to Māori.
TN	Total nitrogen.
TP	Total phosphorus.
TSS	Total suspended sediment.

Abbreviation/ term	Explanation
Turbidity	Is a measure of the cloudiness or haziness in a liquid caused by light scattering by suspended particulate matter. An increase in turbidity results in a corresponding decrease in water clarity. High turbidity may be from an increase in phytoplankton (algae) or an increase in suspended sediments.
Value	Means: a) any national value, and b) includes any value in relation to estuaries, that is not a national value, which a regional council identifies as appropriate for regional or local circumstances (including any use value). Values are intrinsic qualities, uses or potential uses associated with estuaries. They are qualities or uses that people and communities appreciate about estuaries and wish to see recognised in their on-going management. Intrinsic qualities include ecosystem health, and natural form and character.
Visual clarity	Visual clarity is the maximum distance at which an object (typically a black disk) can be seen horizontally through the water column using an underwater viewing apparatus.

7 References

Bolton-Ritchie, L., Lawton, R. (2017) Summary of Regional Council metadata Collected to inform the Coastal Ecosystem module of LAWA: 39.

Cornelisen, C., Zaiko, A., Hewitt, J., Berthelsen, A., McBride, G., Awatere, S., Sinner, J., Banks, J., Hudson, N. (2017) Managing Upstream: Estuaries State and Values: Stage 1A report. *Client report* 2017221HN, prepared for the Ministry for the Environment: 98.

Dudley, B., Zeldis, J., Burge, O. (2017) New Zealand Coastal Water Quality Assessment. Prepared by NIWA for the Ministry for the Environment. *NIWA client report* No: 2016093CH.

Robertson, B.M., Gillespie, P.A., Asher, R.A., Frisk, S., Keeley, N.B., Hopkins, G.A., Thompson, S.J., Tuckey, B.J. (2002) Estuarine Environmental Assessment and Monitoring: A National Protocol. Part A. Development, Part B. Appendices, and Part C. Application. Prepared for supporting Councils and the Ministry for the Environment, *Sustainable Management Fund Contract,* No. 5096. Part A: 93. Part B: 159. Part C. 40. plus field sheets.

Robertson, B.M., Stevens, L., Robertson, B., Zeldis, J., Green, M., Madarasz-Smith, A., Plew, D., Storey, R., Hume, T., Oliver, M. (2016a) NZ Estuary Trophic Index Screening Tool 1. Determining eutrophication susceptibility using physical and nutrient load data. Prepared for Envirolink Tools Project: Estuarine Trophic Index, *MBIE/NIWA Contract,* No: C01X1420: 47.

Robertson, B.M., Stevens, L., Robertson, B., Zeldis, J., Green, M., Madarasz-Smith, A., Plew, D., Storey, R., Oliver, M. (2016b) NZ Estuary Trophic Index Screening Tool 2. Determining Monitoring Indicators and Assessing Estuary Trophic State. Prepared for Envirolink Tools Project: Estuarine Trophic Index, *MBIE/NIWA Contract,* No: C01X1420: 68.

Townsend, M., Lohrer, A.M. (in prep.) ANZECC Guidance for Estuary Sedimentation. *NIWA Client Report* No: HAM2015-096. Prepared for Ministry for the Environment.

Appendix A Factsheets on methods used for monitoring of attributes

Water quality attributes

Water TN, TP

Nutrients fuel aquatic primary producers such as benthic microalgae, photosynthetic bacteria, phytoplankton, macroalgae and aquatic vascular plants (Duarte 1995). Excess nutrient loading can lead to enhanced growth of primary producers which may degrade estuarine habitats, water quality, and be toxic to humans and other consumers (Paerl et al. 2014). Measurements of water total nitrogen (TN) and total phosphorus (TP) respond to upstream loading of nutrients, although their importance in estuaries will depend on benthic fauna and flora cycling of nutrients and inputs from other sources (e.g., ocean) (Bricker et al. 2003). Estuary typology will also influence the impact of nutrients on estuaries. Measurements of water TN and TP will be more closely linked to upstream pressures in shallow, poorly flushed estuaries than highly flushed ones, or those that have a strong oceanic or within estuary inputs (Robertson et al. 2016). As nutrient concentrations are highly temporally variable, measurements require high-frequency sampling that can be expensive. Nutrient concentrations are also highly spatially variable, requiring high spatial resolution of sampling. Further, links between water column nutrient concentrations and ecosystem health status have not been demonstrated. Due to the high spatiotemporal variability nutrient concentrations may have limited use as an attribute. Modelled nutrient loads and/or potential concentrations which account for this variability may provide a viable alternative measure.

Water Chlorophyll a

Water chlorophyll a concentration is a proxy for phytoplankton biomass. Phytoplankton are a food source for many estuarine species and play an important role in estuarine and coastal nutrient cycling (Cloern et al. 2014). Chlorophyll a can increase with nutrient loading from upstream sources due the proliferation of phytoplankton; high concentrations of chlorophyll a is a sign of eutrophication (Boyer et al. 2009). However, it is difficult the separate the response of chlorophyll a to different stressors. For example, as phytoplankton and other primary producers require light to photosynthesis, chlorophyll a can respond negatively to increased sediment loading due to lower light levels (Cloern et al. 2014). Further, chlorophyll a concentration vary spatially and temporally variable within and between estuaries. Smaller estuaries with reduced tidal flushing may be more susceptible to measurable changes in chlorophyll a (i.e., phytoplankton biomass) compared to larger frequently flushed estuaries. Links between chlorophyll a concentrations and ecosystem health status have not been demonstrated.

Water clarity/ Turbidity/TSS

Both water clarity and turbidity can be used as proxies for total suspended solids (TSS). TSS includes sediment and organic material including phytoplankton in the water column. TSS increases with increased sediment loading during rain and flood events, and also during resuspension of mud and sediments within the estuary which can occur during wind/wave events (Dyer 1997, Lawson et al. 2007). TSS also increases with increased phytoplankton biomass. Measures of TSS are highly variable even on the shortest of time scales and can respond to factors other than upstream pressures (e.g., resuspension) (Harris et al. 2015). Shallow, exposed estuaries and deeper estuaries with strong currents will be more susceptible to within estuary generated TSS (i.e., resuspension of seabed sediments). Poorly flushed and smaller estuaries will have a reduced capacity to dilute and assimilate TSS inputs. Estuaries with high sediment loading from their catchment and riverine inputs can have persistently high levels of suspended sediment, despite daily exchange and export to the coastal marine environment.

Increased TSS reduces light penetration, which can affect primary production and in turn food availability for suspension feeders (Cloern et al. 2014); increased turbidity can affect "sight" predators such as some birds and fish. Suspended sediments can be deleterious for marine benthic fauna when concentrations are high enough to clog respiratory or feeding structures. Many species are known to be highly sensitive to suspended-sediment stress (Hewitt et al. 2001, Norkko et al. 2006).

Measures of TSS typically involve collecting and filtering water samples and weighing solids that remain on the filter. An alternative to directly measuring TSS is measuring water clarity which can be easy and cost effective (e.g., using a Sachi disk). However, spot sampling for TSS or water clarity is unlikely to account for the high temporal variation in TSS. Continuous measures of turbidity using in situ sensors can be used to get around this limitation, and can be calibrated against water samples of TSS to covert turbidity values to estimated TSS. However, deployments of continuous water quality sensors can be relatively expensive and have issues in regards to comparability between sensors (Dudley et al. 2017). A key gap that needs to be addressed is how measures of TSS/turbidity/water clarity can be meaningfully applied to relevant thresholds.

Sampling design (applies to all)

National standards / guidelines and consistency: No national standards. Sampling varies across the country (i.e., site extent, number of replicates, frequency and sampling time). National Environmental Monitoring Standards (NEMS) is currently developing guidelines for water quality sampling and analyses with an expected publish date of 1 December 2017.

Potential bottlenecks: Inappropriate sampling design, in terms of spatial extent and number of replicates, selection of non-representative sampling sites.

Opportunities: Water sampling can be aligned (e.g., chlorophyll a, turbidity, clarity, TSS) to reduce sampling effort and improve interpretation.

Caveats and recommendations: The following recommendations are taken directly from the New Zealand Coastal Water Quality Assessment (Dudley et al. 2017).

- Sites should be split proportionally across hydrosystem types/regions.
- Sites should be replicated sufficiently with respect to environmental classes of catchment land use.
- Nutrients affecting coastal hydrosystems should be assessed by monitoring water quality in terminal river reaches, within estuaries and on their adjacent coasts.

Sampling procedures (applies to all)

National standards / guidelines and consistency: Sample collection methods vary in regards to collection platform (boat, wading etc.), collection depth (surface grab, integrated tube taken from the top 15 m, etc.), and collection timing (season, tidal state, time of day, etc.), instrumentation etc.

No national standards, however National Environmental Monitoring Standards (NEMS) is currently developing guidelines for water quality sampling and analyses with an expected publish date of 1 December 2017. *Water clarity* - Typically measured using either a black or Sachi disk. Secchi disk is recommended. **Potential bottlenecks:** Inconsistent sampling methodology in combination with high spatiotemporal variability limits comparability of datasets. **Opportunities:** ▪ Standardising collection methods will improve comparability of datasets. ▪ Standardising of sampling equipment/loggers and calibration protocols will improve comparability of datasets. ▪ Water sampling can be aligned (e.g., nutrients, chlorophyll a, turbidity, clarity, TSS) to reduce sampling effort and improve interpretation. **Caveats and recommendations**: ▪ There should be unified use of the water quality attributes/variables of interest. ▪ Adoption of the National Environmental Monitoring Standards (NEMS).
Laboratory analyses (Applies to nutrients and chlorophyll a)
National standards / guidelines and consistency: Laboratory analysis methods to date have varied considerably and have unknown comparability. No national standards, however National Environmental Monitoring Standards (NEMS) is currently developing guidelines for water quality sampling and analyses with an expected publish date of 1 December 2017. *Nutrients* ▪ Councils typically measure TN using either alkaline persulfate digestion method or sulphuric acid digestion procedure to measure total Kjeldahl nitrogen (TKN). These measures are suggested to give comparable results other than for samples with high suspended solid loads. Relationships between the two methods need to be researched for marine waters. Other methods of TN are less commonly used are unlikely to give comparable results. ▪ Only TP measured by the persulfate digestion method with unfiltered samples were retained for analysis in the Dudley report. *Chlorophyll a* ▪ Typically measured using acetone pigment extraction, spectrofluorometric measurement or in situ and laboratory fluorometry. Acetone pigment extraction, spectrofluorometric measurement is recommended. **Potential bottlenecks:** Variable analytical methods with limited comparability.

<u>Opportunities:</u> Standardise laboratory analysis and/or develop methods to enable comparison of multiple method types.

<u>Caveats and recommendations</u>:

- Reporting uncensored data values by laboratories is strongly recommended (Dudley et al. 2017).
- Adoption of the National Environmental Monitoring Standards (NEMS).

Computational approaches and metrics derived (Applies to nutrients and TSS)

<u>National standards / guidelines and consistency:</u> No national standards.

<u>Potential bottlenecks:</u>

Nutrients

Water column nutrient concentrations do not necessarily reflect the quantity of nutrients available to primary producers. For example, primary producers such as phytoplankton or macroalgae may reduce nutrient concentrations to negligible levels to fuel algal growth.

TSS

Spot sampling for TSS is likely to be negatively impacted by the high temporal variation in suspended sediment loads. For example, suspended sediment loads may be highest during storm events when spot sampling is impractical or unsafe.

<u>Opportunities:</u>

Nutrients

Measures of nutrient loads have greater biological relevance and application in thresholds than nutrient concentrations. Nutrients affecting coastal hydrosystems should be assessed by monitoring water quality in terminal river reaches, within estuaries and on their adjacent coasts. GIS based tools such as CLUES-Estuary can be used to mix loads entering estuaries to estimate the nutrients available to the primary producers prior to uptake (within estuaries), or the nutrient loading.

TSS

Combining spot sampling for TSS with continuous measures of turbidity using turbidity sensors and/or models can be used to better account for temporal variation in TSS and improve estimates sediment loads and threshold setting.

<u>Caveats and recommendations</u>:

- There should be unified use of NEMS protocols with regard to metadata collection, reporting of measurement uncertainty and quality coding (Dudley et al. 2017).
- Measures of water nutrient concentrations should generally be integrated into measures/estimates of nutrient loads which have greater biological relevance and application for threshold setting (e.g., through the use of models such as CLUES-Estuary).

▪ Measures of water TSS loads should account for the high temporal variation in values, which has important consequences for threshold setting (e.g., through the paired use of continuous turbidity loggers).
Thresholds (applies to all)
No national standards. ANZECC guidelines or a derivative thereof is used (WRC) or being investigated (AC) to determine thresholds for nutrient concentrations. HRC set nutrient concentration thresholds based on their OnePlan which takes into account estuarine typology. National Environmental Monitoring Standards (NEMS) is currently developing guidelines for water quality sampling and analyses with an expected publish date of 1 December 2017. The following recommendations are taken from the report "New Zealand Coastal Water Quality Assessment" (Dudley et al. 2017). ▪ The setting of water quality thresholds should account for characteristics of different hydrosystem types – some hydrosystem types are more sensitive to stressors than others. ▪ We would not recommend using the current water quality dataset for threshold setting using a percentile-based approach because 1) the dataset is not representative of water quality conditions in New Zealand coastal hydrosystems nationally, for the reasons laid out in Section 5, and 2) we currently do not fully understand how levels for each water quality variable relate to values (such as ecosystem health). ▪ We recommend that thresholds for water quality and contaminant loads are set by comparing hydrosystem water quality with scores of ecosystem health and other values. ▪ We recommend further development of relationships between contaminant loading rates, water quality, and hydrosystem ecological health to inform water quality threshold setting. ▪ An integrated index of hydrosystem ecological health should be included in future state and trend analysis to facilitate setting of water quality thresholds (i.e., boundaries between bands of environmental state) and increase the utility of monitoring.
Emerging and prospective future methods
The Dudley et al. (2017) report recommends the following three variables (Light availability, CDOM, Munsell colour) as supporting variables by NEMS (*in prep*) in addition to those optical variables analysed for state and trends in this report (CLAR, SS and TURB). Currently these variables do not appear to be a significant component of council monitoring programs (Dudley et al. 2017). *Light availability*

- Direct biological relevance as light drives photosynthesis. Can be measured using continuous loggers. Captures changes in suspended sediment loads. Can be applied to generate light based thresholds with high biological relevance.

CDOM

- Coloured dissolved organic matter (CDOM) is a useful index of freshwater content of water that correlates inversely with salinity. Both salinity and CDOM have useful application in remote sensing of estuarine and coastal water quality as well as relationship to water values. CDOM also provides a measure of organic carbon transport from land to the ocean.

Munsell Colour

- Munsell colour is a valuable observation on water optical character that can be useful in QA of water quality.

References

Boyer, J. N., C. R. Kelble, P. B. Ortner, and D. T. Rudnick. 2009. Phytoplankton bloom status: Chlorophyll a biomass as an indicator of water quality condition in the southern estuaries of Florida, USA. Ecological Indicators **9**:S56-S67.

Bricker, S. B., J. G. Ferreira, and T. Simas. 2003. An integrated methodology for assessment of estuarine trophic status. Ecological Modelling **169**:39-60.

Cloern, J. E., S. Foster, and A. Kleckner. 2014. Phytoplankton primary production in the world's estuarine-coastal ecosystems. Biogeosciences **11**:2477-2501.

Duarte, C. M. 1995. Submerged aquatic vegetation in relation to different nutrient regimes. Ophelia **41**:87-112.

Dudley, B., J. Zeldis, and B. O. 2017. New Zealand Coastal Water Quality Assessment. Prepared by NIWA for the Ministry for the Environment. NIWA client report No: 2016093CH.

Dyer, K. R. 1997. Estuaries: A Physical Introduction, 2nd Edition, Wiley Press: 210.

Harris, R. J., C. A. Pilditch, J. E. Hewitt, A. M. Lohrer, C. Van Colen, M. Townsend, and S. F. Thrush. 2015. Biotic interactions influence sediment erodibility on wave-exposed sandflats. Marine Ecology Progress Series **523**:15-30.

Hewitt, J. E., S. Hatton, K. Safi, and R. Craggs. 2001. Effects of suspended sediment level on suspension-feeding shellfish in the Whitford embayment. Report prepared by NIWA for Auckland Regional Council. Auckland Council Technical Publication 2001/159.

Lawson, S. E., P. L. Wiberg, K. J. McGlathery, and D. C. Fugate. 2007. Wind-driven sediment suspension controls light availability in a shallow coastal lagoon. Estuaries and Coasts **30**:102-112.

Norkko, A., J. E. Hewitt, S. F. Thrush, and G. A. Funnell. 2006. Conditional outcomes of facilitation by a habitat forming subtidal bivalve. Ecology **87**:226-234.

Paerl, H. W., N. S. Hall, B. L. Peierls, and K. L. Rossignol. 2014. Evolving paradigms and challenges in estuarine and coastal eutrophication dynamics in a culturally and climatically stressed world. Estuaries and Coasts **37**:243-258.

Robertson, B., M, L. Stevens, B. Robertson, J. Zeldis, M. Green, A. Madarasz-Smith, D. Plew, R. Storey, and M. Oliver. 2016. NZ Estuary Trophic Index Screening Tool 2. Determining Monitoring Indicators and Assessing Estuary Trophic State. Prepared for Envirolink Tools Project: Estuarine Trophic Index, MBIE/NIWA Contract No: C01X1420. 68p.

National Environmental Monitoring Standards (NEMS). http://www.nems.org.nz/

Water faecal indicator bacteria

Faecal indicator bacteria (FIB) are present in the gut of animals and excreted with faeces that provide an indication of faecal contamination of water and are used to assess the risk of humans developing illness associated with contact with contaminated water. New Zealand councils monitor *Escherichia coli* in freshwater and *Enterococcus* spp. in marine waters to assess disease risk. *Escherichia* coli concentrations are indicative of the risk of infection (and possibly illness) from pathogens such as *Campylobacter* (Ministry for the Environment 2017) from contact with waters contaminated with faecal material (McBride et al. 1998). Enterococci are used for monitoring the risk associated with contact recreation at marine beaches as McBride et al. (1998) found that enterococci had the stronger correlation with disease risk. Faecal coliforms are less specific to humans than *E. coli* and enterococci but are considered more suitable for general assessments of faecal contamination in shellfish gathering water (Ministry for the Environment 2003). As summarized in Milne et al. (2017), it remains unclear from existing national guidelines as to whether *E. coli* or enterococci should be tested on samples from estuarine waters. Some councils test for *E. coli*, some for enterococci, and some for both indicator bacteria. Advice on the most appropriate indicator to use in estuaries (brackish waters) is a focal point of a recently supported Envirolink Tools project aimed at addressing coastal water quality guidelines (McBride 2016).
Sampling design
<u>National standards / guidelines and consistency</u>: In New Zealand, monitoring of microbial contamination is common within state of the environment (SOE) water quality monitoring programmes and can be used to illustrate where environmental management has been effective. Microbial monitoring of freshwater and marine recreational bathing sites is also carried out during the summer bathing season for routine weekly surveillance and longer term grading purposes. Standards for microbial contamination are also routinely put in place for consented point source discharges (e.g., sewage outfalls). Sampling times and period to measure bathing water quality at beaches are set out at http://www.mfe.govt.nz/publications/international-environmental-agreements/microbiological-water-quality-guidelines-marine#notehi. To manage bathing waters, samples are typically collected at popular bathing beaches, and results are used to grade a beach according to MfE/MoH 2003 guidelines. Decisions around bathing closures are based on microbial monitoring data and, in some jurisdictions, on recent rainfall and/or adjacent river flows (circumventing the delay in microbial analysis, typically one day, so that swimming advisories are a day too late). **<u>Potential bottlenecks:</u>** Inappropriate sampling design, in terms of spatial extent and number of replicates, selection of non-representative sampling sites. Samples also tend to be biased toward good weather conditions, rather than being collected for example during periods of high rainfall. As outlined in Milne et al. (2017), there is also a lack of consensus around the state measure, statistic and minimum sample size to report and a universally applied approach to determining a meaningful improvement or decline in water quality. **<u>Opportunities:</u>** FIB data could be used to ground-truth predictive, operational models for estimating current FIB concentrations until more rapid techniques are available. Milne et al. (2017) outlines a

number of measures needed to improve accuracy, robustness, and meaningfulness of recreational water quality monitoring. **Caveats and recommendations**: Several factors need to be considered when collecting water for FIB tests, and when interpreting results. As outlined in Green and Cornelisen (2013), risk of faecal contamination varies according to surrounding catchments and land use and the hydrological characteristics of the coastal water body (e.g., flushing). Wave action, climate and water depth also influence FIB concentrations as the bacteria are known to persist in sediments and beach sands and may spike without recent rainfall. Bacteria and viruses are also more prevalent in turbid waters where microbes attach to particles that prolong survival due to solar shading and extend microbe transport distance. As a result, there are a number of parameters that may influence levels of microbial contamination (elevated FIB), including rainfall, solar radiation, tidal state, water clarity and suspended sediments (or turbidity), light penetration, salinity and water temperature. Due to high variability, modelled estimates for estuary FIB concentrations in response to upstream loading based on land uses and varying conditions may be required to develop an FIB attribute, whereas measured concentrations could be used for supporting the attribute and as a state variable.
Sampling procedures
National standards / guidelines and consistency: Sampling protocols to measure recreational water quality are set out at http://www.mfe.govt.nz/publications/international-environmental-agreements/microbiological-water-quality-guidelines-marine#notehi. **Potential bottlenecks:** Bottlenecks include those as outlined under sampling design and relate to the timing and conditions when samples are collected, and insufficient spatial and temporal replication of samples.
Laboratory analyses
National standards / guidelines and consistency: For evaluating faecal coliform bacteria concentrations, Membrane filtration (APHA 9222D) and Multiple tube (APHA 9221E) measurement procedures are being used. Both procedures are presumed to give comparable results. For enterococci concentrations, Multiple tube (APHA 9230B), Membrane filtration (APHA 9230C) and Fluorogenic Substrate Enterococcus Test 'Enterolert' (APHA 9230D) are applied. All procedures are presumed to give comparable results (Dudley et al. 2017). Limited investigation has indicated that further work in this area is required. For example, some agencies use different methods of analysis for samples derived from State of Environment monitoring programmes than for samples collected for recreational water quality monitoring programmes, but others use a single method for samples collected for either programme. **Potential bottlenecks:** Currently culture based methods that require at least 24 hours incubation are used to measure *E. coli* and enterococci concentrations. The incubation period results in warnings of faecal contamination after events such as sewerage spills or storms until beaches can be shown to be non-contaminated. **Opportunities:** A faster method of assessing FIB would provide more timely warnings and improve compliance with the warnings. Predictive models would avoid the time-delay problem. Also, identifying host source of faecal contamination would provide better understanding of the relative risk associated with the faecal contamination.

Computational approaches and metrics derived
National standards / guidelines and consistency: A coastal beach's suitability for recreation is assessed from a Microbiological Assessment Category (MAC) based on up to five years of enterococci test results (for marine beaches) (Table 1-1) and a Sanitary Inspection Category (SIC) based on a beach's risk of human faecal pollution. Beaches are given a Suitability for Recreation Grade (SFRG) ranging from Very Good, Good, Fair, Poor or Very Poor based on a combination of the MAC and SIC results. Beaches graded Good, Fair or Poor have the potential to be affected by faecal contamination and must be tested routinely (e.g., weekly) for enterococci concentrations. There are also Guidelines for follow-up day-by-day surveillance of beaches when results exceed acceptable levels of bacteria concentrations (see Table 2 below). **Caveats and recommendations**: Natural patchiness in the distribution of faecal indicator bacteria can impede the ability to identify trends over time. For instance, enterococci concentrations in coastal waters have been shown to vary by 60% on average and by as much as 700% between samples that are collected only minutes apart (Boehm, 2007).
Thresholds
The Microbiological Water Quality Guidelines for Marine and Freshwater Recreational Areas set thresholds for water quality (MfE/MoH 2003, Tables 1 and 2).

Table 1: MfE/MoH Microbiological Assessment Category (MAC) definitions for marine waters for site grading.

Grade	Standard
A	Sample 95 percentile ≤ 40 enterococci per 100 mL
B	Sample 95 percentile 41–200 enterococci per 100 mL
C	Sample 95 percentile 201–500 enterococci per 100 mL
D	Sample 95 percentile > 500 enterococci per 100 mL

Table 2: MfE/MoH trigger points for marine waters. Enterococci concentrations are typically expressed using the most probable number (MPN) method for cultivable bacteria.

	Level of action	Standard
Green	Routine surveillance	No single sample > than 140 enterococci per 100 mL
Amber	Alert	Single sample > 140 enterococci per 100 mL
Red	Action	Two consecutive samples > 280 enterococci per 100 mL

The trigger points comprise a three-tier system analogous to traffic lights:

Green – highly likely to be uncontaminated: "suitable for bathing", but requiring water managers to continue surveillance (e.g., weekly testing for enterococci). **Amber** – potentially contaminated: "potentially unsuitable for bathing", requiring water managers to investigate the suitability for recreation, increase testing for enterococci to daily, and identify sources of contamination. **Red** – highly likely to be contaminated: "highly likely to be unsuitable for bathing", requiring urgent action from water managers, including daily testing for enterococci and public warnings and identify sources of contamination.
Emerging and prospective future methods As summarised in Green and Cornelisen (2016), emerging technologies for monitoring FIB may replace or complement culture based tests for FIB as they become validated. Tests for FIB that are faster than the current culture based tests will address the current challenge around delayed results; typically results using standard culture methods cannot be produced for at least 24 hours following sample collection. The U.S. EPA recently approved some rapid quantitative polymerase chain reaction (qPCR) based methods for FIB (for example, the Bacteriological Analytical Manual, Method 1609 for enterococci (Anonymous 2013)). Technological advances in instrumentation using methods that do not rely on cultivation to detect FIB also offer the potential to evaluate contamination in real-time and to be used as in-situ sensors for remote monitoring (Lebaron et al. 2005, Ryzinska-Paier et al. 2014). Microbial Source Tracing (MST) uses DNA-based markers that can be quantitative and identify host specific species of bacteria and viruses allowing the identification of the source of the faecal contamination. Currently, the technology is used to prioritise and solve contamination problems. In the coming years, there is likely to be a number of source-specific markers that may be implemented within a monitoring programme to inform risk management (e.g., closing a beach due to presence of faecal contamination from humans as opposed to seabirds). An Envirolink Tools project organized through the Coastal Special Interest Group included a review and trial of MST markers for use in New Zealand coastal waters. Water samples were collected on four occasions across 53 sites and then analysed for FIB and MST markers (Cornelisen et al. 2012). A Bacteroidales marker specific to contamination from ruminant animals (cows and sheep) showed the greatest promise as a tool to inform water quality monitoring programmes. With advancing molecular technologies, there is also the potential to directly measure pathogens as opposed to indicators of pathogen presence. Milne et al. (2017) suggest monitoring *Campylobacter* spp. *Cryptosporidium*, norovirus and adenovirus, at problematic or 'high risk' sites as a good starting point.
References
Anonymous (2013) Method 1609: Enterococci in Water by TaqMan® Quantitative Polymerase Chain Reaction (qPCR) with Internal Amplification Control (IAC) Assay Washington, D. C. , EPA United States Environmental Protection Agency**:** 72. Cornelisen, C.D., Kirs, M., Gilpin, B., Scholes, P. (2012) Microbial Source Tracking (MST) tools for water quality monitoring. Prepared for Regional Councils and the Special Interest Group. Cawthron Report No. 2047. 28 p. plus appendices. Dudley, B., Zeldis, J., Burge. O. (2017) New Zealand Coastal Water Quality Assessment. Prepared for Ministry for the Environment. NIWA CLIENT REPORT No: 2016093CH.

Green, M.O. and Cornelisen, C.D. (2016) Marine Water Quality Standards for the Waikato Region – Literature Review. Report WRC1507–1, Streamlined Environmental, Hamilton, 123 pp.

Lebaron, P., Henry, A., Lepeuple, A-S., Pena, G and Servais, P (2005). An operational method for the real-time monitoring of E. coli numbers in bathing waters. Marine Pollution Bulletin, 50 p652-659

McBride, G. B., Salmond, C. E., Bandaranayake D. R., Turner S. J., Lewis G. D.,Till, D. G. (1998) Health effects of marine bathing in New Zealand. *International Journal of Environmental Health Research* 8: 173-189.

McBride, G. (2016). Marine contact recreation water and shellfish guidelines Envirolink Tools project 2016/17. Prepared for the Ministry for Business, Innovation and Excellence on behalf of the regional sector's Coastal Special Interest Group. Contract C01X1610.

Milne, J., Madarasz-Smith, A., Davie, T. (2017) Recreational water quality monitoring and reporting; A position paper prepared for the NZ regional sector. 28 p.

Ministry for the Environment (2003) Microbiological Water Quality Guidelines for Marine and Freshwater Recreational Areas. Ministry for the Environment, Wellington, New Zealand.

Ministry for the Environment (2017) National Policy Statement for Freshwater Management 2014 Updated August 2017 to incorporate amendments from the National Policy Statement for Freshwater Amendment Order 2017**:** 47.

Ryzinska-Paier, G., Lendenfeld, T., Correa, K., Stadler, P., Blaschke, A.P., Mach, R.L.,Stadler, H., Kirschner, A.K., Farnleitner, A.H., 2014. A sensitive and robust method for automated on-line monitoring of enzymatic activities in water and water resources. Water Sci. Technol. 69, 1349e1358. http://dx.doi.org/10.2166/wst.2014.032

Macroalgae

Many methods have been developed in the world to assess estuary eutrophication and allow regulatory authorities to meet statutory requirements (e.g., to monitor and protect estuaries from degradation). These methods demonstrate that the eutrophication gradient is well understood and that the immediate biological response is increased primary production reflected as increased chlorophyll a and/or macroalgal abundance, which is often accompanied by secondary symptoms within both the water column and sediments. Primary symptoms (e.g., macroalgae outbreaks) are considered to exhibit unambiguous responses to eutrophication. Supporting indicators can have variable and/or ambiguous relationships with eutrophication but are useful in its measurement. As a result, most methods include both primary symptoms and supporting indicators to provide the best possible evaluation of the nutrient related quality of the water body (Borja et al. 2012, Devlin et al. 2011, Sutula 2011).

Opportunistic macroalgae are species that survive well in conditions in which other species often struggle to survive or compete (Borum and Sand-Jensen 1996). Blooms in NZ estuaries principally contain species of green algae *Ulva* (this includes taxa formerly known as *Enteromorpha*) and *Cladophora*, red algae *Gracilaria*, and brown algae (e.g., *Ectocarpus, Pilayella, Bachelotia*). These bloom-forming species are a natural component of intertidal ecosystems (Adams 1994), but they only grow to bloom proportions when nutrient levels are elevated (Sutula et al. 2011) and sufficient light reaches the bed of the estuary (or the water column where macroalgae are suspended). As a consequence, they generally only reach nuisance conditions in shallow estuaries, or the margins of deeper estuaries. The macroalgal response to nutrient loads generally increases with water residence times (Painting et al. 2007), either of the whole estuary (as is often the case for many NZ short residence time estuaries), or part of the estuary (e.g., a poorly flushed upper estuary arm where nutrient-rich muds accumulate), or in 'backwaters' where drifting suspended macroalgae can accumulate (e.g., Avon-Heathcote Estuary: Bolton-Ritchie and Main 2005). There is some evidence this response may also be significantly attenuated by the presence of fringing saltmarsh, due to reductions in nutrient loading through processes such as denitrification (Valiela et al. 1997). Other factors that can influence the expression of macroalgal growth are the presence of suitable attachment strata, and physical and hydrodynamic conditions e.g., temperature (desiccation), fetch (wind driven waves), currents (scouring) e.g., Hawes and Smith (1995).

Blooms of rapidly growing macroalgae can have deleterious effects on intertidal and shallow subtidal communities, and cause an undesirable imbalance with effects such as: blanketing of the surface causing a hostile physico-chemical environment in the underlying sediment, sulphide poisoning of infaunal species, anoxic gradient at the water sediment interface, effects on birds including changes in the feeding behaviour of waders, smothering of seagrass beds - (Duarte 1995, Taylor et al. 1995, Valiella et al. 1997, Sutula et al. 2012), excessive algal growths, or rafts of floating or detached weed causing interference with water users, aesthetic effects such as nuisance odours, or deposition in bathing waters. Where excessive macroalgae cause extreme sediment anoxia (measured by redox potential) there is an accompanying exclusion of normal communities of benthic macrofauna (e.g., Grizzle and Penniman 1991); increased production of sulphides which can be toxic to rooted macrophytes (Lamers et al. 2013, Holmer and Bondgaard 2001, Viaroli et al. 2008, Geurts et al. 2009, Green et al. 2014), and release of dissolved phosphorus and ammonium that exacerbate eutrophication (e.g., Søndergaard et al. 2003).

Sampling design
<u>**National standards / guidelines and consistency:**</u> The WFD-UKTAG (Water Framework Directive – United Kingdom Technical Advisory Group, 2014) approach for opportunistic macroalgal condition is a relatively comprehensive rating tool that is currently used on NZ estuaries and is recommended for use in the ETI (Robertson et al. 2016b). It is supported by extensive studies of the macroalgal condition in relation to ecological responses in a wide range of estuaries. The Opportunistic Macroalgal Blooming Tool (OMBT) is a comprehensive 5 part multimetric index that incorporates species composition, macroalgal cover, biomass, and entrainment within sediment to calculate an ecological quality rating (EQR). It is currently used in broad scale assessment of estuary condition by many regional councils in NZ. The OMBT has been developed to classify data over the maximum growing season so sampling should target the peak bloom in spring-summer (Oct-March), although peak timing may vary among water bodies, therefore local knowledge is required to identify the maximum growth period. Sampling is not recommended outside the spring-summer period due to seasonal variations that could affect the outcome of the tool and possibly lead to misclassification; e.g., blooms may become disrupted by stormy autumn weather and often die back in winter. Sampling is best carried out during spring low tides in order to access the maximum area of the Available Intertidal Habitat (AIH). <u>**Potential bottlenecks:**</u> The OMBT has been developed with thresholds to define ecological quality status based on extensive European data. The NZ macroalgal data assessed to date are largely consistent with the established UK-WFD thresholds, but the threshold for significant sediment related impacts appears to occur at a lower macroalgal biomass in NZ than in the UK-WFD (Wriggle, unpublished). Because the OMBT is designed to allow for specific changes such as this to be incorporated, NZ specific thresholds can easily be incorporated. However, a full assessment of available data is needed to apply this in a nationally consistent manner. Currently, available data are scattered throughout individual reports and there has been no collation of national data. Further, the available data in NZ are currently dominated by South Island and lower North Island estuaries and include very few sites in northern estuaries where mangroves are present. The relationship between nutrient loads and macroalgal response may be significantly different in these estuaries and needs to be validated to ensure thresholds for these estuary types are appropriate. Like most sampling there is also potential for variation in the application of the sampling design, particularly in terms of spatial extent and number of replicates, ensuring representative sampling sites are selected, and that criteria used to set thresholds of impact reflect the entire gradient of response to nutrient loads (low/pristine to high/degraded). <u>**Opportunities:**</u> The development of integrated GIS based mapping outputs and calculators would simplify reporting and national consistency. <u>**Caveats and recommendations**</u>: The UK recognize the specialist skillset needed to maintain consistency in macroalgal monitoring using the OMBT, and undertake this work using a specialist provider at a national level, rather than using multiple regional providers. Such an approach should be considered in NZ.

Sampling procedures
<u>National standards / guidelines and consistency</u>: No national standards. Macroalgae was not included as a primary symptom of eutrophication in the NEMP so no specific methods were developed for its enumeration. While it has commonly been recorded where it is a dominant surface cover, NEMP spatial mapping does not include the measures of estuary wide percentage cover, biomass or entrainment that are required by the OMBT. Broad scale spatial mapping described in the NEMP requires updating to reflect subsequent advances in the protocol, and in particular the application of complimentary fine scale measures to delineate substrate boundaries and validate substrate classifications. **<u>Potential bottlenecks:</u>** Adequate training is required to consistently assess and enumerate broad scale macroalgal condition. Field sampling requires the ability to consistently define representative patches of macroalgal cover and biomass, and balance replication needs with practical considerations in terms of sampling within the limited tide window. **<u>Opportunities:</u>** Drones provide a rapid way to quickly assess percentage cover and determine where ground truthing should be concentrated at a local scale. At this point it is difficult to envisage drones being used to assess biomass (as opposed to cover) or entrainment. Remote sensing tools such as infrared cameras may enable automated mapping of macroalgal cover, species composition, and possibly estimates of density. **<u>Caveats and recommendations</u>**: Development of standardized methods for the field measurement of biomass, percentage cover, and entrainment are required to ensure national consistency. Zones of extreme sediment degradation, called "Gross Nuisance Areas (GNAs)", are currently used in the ETI as an indicator of excessive opportunistic macroalgae (including epiphytes) that are associated with anoxic sediment (Robertson and Stevens 2013). Such findings are supported by widespread monitoring of NZ shallow estuaries which indicate that excessive macroalgal cover in poorly flushed parts of these estuaries can result in GNAs (i.e., combined conditions of high mud content, surface sediment anoxia, elevated organic matter and nutrient concentrations, an imbalanced benthic invertebrate community and seagrass die off (Robertson and Stevens 2013). Similar GNAs occur in shallow coastal lagoons or ICOLLs where conditions are not too turbid e.g., Waituna Lagoon. As a consequence, the use of macroalgal abundance as a trophic state indicator must be used alongside other supporting indicators, such as mud content and RPD (e.g., Sutula et al. 2012) in order to accurately predict the trophic status of such estuaries. The presence of persistent and extensive areas of GNAs in estuaries, however, provides a clear signal that the assimilative capacity of the estuary is being exceeded.
Computational approaches and metrics derived
<u>National standards / guidelines and consistency:</u> The OMBT has been developed with thresholds to define ecological quality status based on extensive European data. This includes a full description of the metrics used and calculations required. These are presented in the ETI (Robertson et al. 2016b) and the OMBT (WFD-UKTAG 2014), and a calculator has been developed to automate calculations including confidence measures (Davy 2009). Called CAPTAIN ('Confidence And Precision Tool Aids aNalysis') it calculates confidence of class (CofC) at three levels: i. metric, ii. survey (single sampling event), and iii. water body over the reporting period (potentially several surveys).

Potential bottlenecks: Absence of a collated national dataset of existing data, potential variability in the assessments undertaken by different providers, and uncertainty or inconsistency in the ground truthing undertaken in different estuaries. **Opportunities:** Annual monitoring at a nationally consistent level will provide a rapid and direct measure of eutrophic expression in NZ estuaries. Combined with nutrient load estimates it will be possible to set nutrient load limits (mean annual average) once a robust relationship has been developed for a range of NZ estuary types. **Caveats and recommendations**: Supporting indicators are an important subcomponent of any measurements and are required to understand the implications of macroalgal expression. In any particular substrate type, sediment oxygenation, organic content and sediment nutrient concentrations are all key measures in assessing the likely impact and duration of macroalgal growths. Ongoing PhD work (Ben Robertson, Uni of Otago) is exploring many of these in more detail.
Thresholds
The ETI proposes thresholds for macroalgal based on OMBT scores. These are in turn based on background data in the OMBT (WDF-UKTAG 2014) and elsewhere as follows: A survey of eight Californian tidal lagoon estuaries (including some ICOLLs) by Sutula et al. (2014) found that macroalgae of 175g.m^{-2}dw (1450g.m^{-2} ww), total organic carbon of 1.1% , and sediment TN of 0.1% were thresholds associated with anoxic conditions near the surface (RPD <1cm). In two Californian estuaries, macroalgal abundances as low as 110-120g.m^{-2} dw (or 840-930g.m^{-2} ww) had significant and rapid negative effects on benthic invertebrate abundance (declining by >67%) and species richness (declining by >19%) within two weeks at most sites (Green et al. 2014). An effects threshold of 500-1000g.m^{-2} ww (wet weight per square metre) was proposed by Scanlan et al. (2007) to avoid effects on benthic macrofauna in estuaries, but the authors emphasised that the proposed thresholds required further validation. McLaughlin et al. (2013) reviewed and tested the biomass thresholds proposed by Scanlan et al. (2007) and considered them reasonable for application to Southern Californian ICOLLs. For example, the review found elimination of surface deposit feeders when macroalgal biomass was in the range of 700-800g.m^{-2} ww. In some situations it is possible for macroalgae to continue growth after being covered by sediment (i.e., entrainment) (WFD UKTAG 2014). A review of monitoring data from 25 typical NZ estuaries (shallow, short residence time estuaries) supports an opportunistic macroalgal biomass "exhaustion" threshold of approximately 1000-2000g.m^{-2} ww above which there was a major shift in the chemistry of the underlying sediment to surface anoxia (RPD at the surface), elevated TOC (>1.5%) and a degraded macrofaunal community (Wriggle Coastal Management database 2009-2014). Such conditions have been used to identify GNAs. Based on the measured detrimental impact on macrofauna in NZ tidal lagoons, it has been estimated that if GNAs cover >15% of the estuary area or >30ha, then estuary ecological condition is seriously impaired.

Waituna Lagoon, a NZ ICOLL, was estimated to have a mean macroalgal biomass of 800-1000g.m^{-2} ww when the lagoon was showing signs of gross eutrophication (RPD at sediment surface) and a degraded seagrass community. At 100-300g.m^{-2} ww the seagrass community was maintained with moderately low levels of stress (Hamilton et al. 2012). Currently, the data supporting a relationship between macroalgae and estuary trophic status in NZ estuaries is limited to a relatively small number of studies, but all confirm adverse impacts to sediment physico-chemistry and biota along similar lines to those found in overseas studies. In order to provide a more robust basis upon which to base the metrics used in the OMBT (WFD-UKTAG 2014) ecological quality rating for macroalgae, it is recommended that the ecological response thresholds for macroalgae be more thoroughly assessed, over all estuary types (but particularly those prone to macroalgal blooms i.e., shallow, intertidal dominated estuaries and ICOLLs). The studies should focus on opportunistic macroalgal effects on biota (e.g., macroinvertebrates, fish, seagrass), and physico-chemical parameters (e.g., sediment redox potential, sulphur, organic carbon, nutrients and bacteria)
Emerging and prospective future methods
Undertake more comprehensive studies to improve our understanding of the relationship between nutrient loads and ecological response in shallow, intertidal dominated estuaries and ICOLLS. In particular, it is recommended that monitoring of the following be undertaken and the data used to establish load response relationships: macroalgal biomass and sediment characteristics (nutrients, organic carbon, sulphur components, redox potential, bacterial composition) and the relationships of these variables with seagrass, mangroves, macroinvertebrates, and fish. Undertake studies quantifying relative nutrient supply (sediment vs water column) and preferential species uptake in order to better understand likely response under various situations.
References
Adams, N.M. 1994. Seaweeds of New Zealand. Christchurch, Canterbury University Press. 360p. Bolton-Richie, L., and Main, M. 2005. Nutrient water quality Avon-Heathcote Estuary/Ihutai: Inputs, concentrations and potential effects. Report no. U05/71 for Environment Canterbury. Borja, A., Basset, A., Bricker, S., Dauvin, J., Elliot, M., Harrison, T., Marques, J., Weisberg, S. and West, R. 2012. Classifying ecological quality and integrity of estuaries. In E. Wolanski & D. McLusky (Eds.), Treatise on Estuarine and Coastal Science: 125-162. Waltham: Academic Press. Borum, J. and Sand-Jensen, K. 1996. Is total primary production in shallow coastal marine waters stimulated by nitrogen loading? Oikos 76:406-410. Davey, A. 2009. Confidence of Class for WFD Marine Plant Tools. WRC report EA7954. 34pp. Devlin, M., Bricker, S. and Painting, S. 2011. Comparison of five methods for assessing impacts of nutrient enrichment using estuarine case studies. Biogeochemistry, 106(2): 177-205. Duarte, C.M. 1995. Submerged aquatic vegetation in relation to different nutrient regimes. Ophelia 41: 87-112. Geurts, J.J.M., Sarneel, J.M., Willers, B.J.C., Roelofs, J.G.M., Verhoeven, J.T.A. and Lamers, L.P.M. 2009. Interacting effects of sulphate pollution, sulphide toxicity and eutrophication on vegetation development in fens: a mesocosm experiment. Environ. Pollut. 157: 2072-2081. Green, L., Sutula, M. and Fong, P. 2014. How much is too much? Identifying benchmarks of adverse effects of macroalgae on the macrofauna in intertidal flats. Ecological Applications 24(2): 300-314. Grizzle, R.E. and Penniman, C.A. 1991. Effects of organic enrichment on estuarine macrofaunal benthos: a comparison of sediment profile imaging from traditional methods. Marine Ecology Progress Series. 74: 249-262.

Hamilton, D.P., Jones, H.F.E., Özkundakci, D., McBride, C., Allan, M.G., Faber, J. and Pilditch, C.A. 2012. Waituna Lagoon Modelling: Developing quantitative assessments to assist with lagoon management University of Waikato. ERI report number: 004.

Hawes, I., Smith, R. 1995. Effect of current velocity on the detachment of thalli of Ulva lactuca (chlorophyta) in a NZ estuary. Journal of phycology, 31: 875-880.

Holmer, M., Bondgaard, E.J. 2001. Photosynthesis and growth response of eelgrass to low oxygen and high sulphide concentrations during hypoxic events. Aquat. Bot. 70: 29-38.

Lamers, L.P.M., Govers, L.L., Janssen, I.C.J.M., Geurts J.J.M., van der Welle, M.E.W. and van Katwijk, M.M. 2013. Sulphide as a soil phytotoxin—a review. Frontiers in Plant Science 4: 268.

McLaughlin, K., Sutula, M., Busse, L., Anderson, S., Crooks, J., Dagit, R., Gibson, D., Johnston, K. and Stratton, L. 2013. A regional survey of the extent and magnitude of eutrophication in Mediterranean estuaries of Southern California, USA. Estuaries and Coasts. dx.doi.org/10.1007/s12237-013-96708 .

Painting, S.J., Devlin, M.J., Malcolm, S.J., Parker, E.R., Mills, D.K., Mills, C. and Winpenny, K. 2007. Assessing the impact of nutrient enrichment in estuaries: susceptibility to eutrophication. Marine Pollution Bulletin, 55(1-6): 74-90.

Robertson, B.M. and Stevens, L.M. (2013). Jacobs River Estuary. Fine scale monitoring of highly eutrophic arms 2012/13. Report prepared for Environment Southland. 29p.

Robertson, B.M., Gillespie, P.A., Asher, R.A., Frisk, S., Keeley, N.B., Hopkins, G.A., Thompson, S.J., Tuckey, B.J. (2002). Estuarine Environmental Assessment and Monitoring: A National Protocol. Part A. Development, Part B. Appendices, and Part C. Application. Prepared for supporting Councils and the Ministry for the Environment, Sustainable Management Fund Contract No. 5096. Part A. 93p. Part B. 159p. Part C. 40p. plus field sheets.

Robertson, B.M., Stevens L.M., Robertson, B.P., Zeldis, J., Green, M., Madarasz-Smith, A., Plew, D., Storey, R., Oliver, M. (2016b). NZ Estuary Trophic Index Screening Tool 2. Determining Monitoring Indicators and Assessing Estuary Trophic State. Prepared for Envirolink Tools Project: Estuarine Trophic Index, MBIE/NIWA Contract No: C01X1420. 68p.

Scanlan, C.M., Foden, J., Wells, E. and Best, M.A. 2007. The monitoring of opportunistic macroalgal blooms for the water framework directive. Marine Pollution Bulletin 55: 162-171.

Søndergaard, M., Jeppesen, E. and Jensen, J.P. 2003. Internal phosphorus loading and the resilience of Danish lakes. Lake Line 23: 17-20.

Sutula, M. 2011. Review of Indicators for Development of Nutrient Numeric Endpoints in California Estuaries. The California Environmental Protection Agency State Water Resources Control Board, Technical Report 646.

Sutula, M., Bailey, H. and Poucher, S. 2012. Science Supporting Dissolved Oxygen Objectives in California Estuaries. Prepared for: The California Environmental Protection Agency State Water Resources Control Board (Agreement Number 07-110-250), (07).

Sutula, M., Green, L., Cicchetti, G., Detenbeck, N. and Fong, P. 2014. Thresholds of Adverse Effects of Macroalgal Abundance and Sediment Organic Matter on Benthic Habitat Quality in Estuarine Intertidal Flats. Estuaries and Coasts. doi:10.1007/s12237-014-9796-3.

Taylor, D.I., Nixon, S.W. Granger, S.L. and Buckley B.A. 1995. Nutrient limitation and the eutrophication of coastal lagoons. Marine Ecology Progress Series 127: 235-44.

Valiela, I., McClelland, J., Hauxwell, J., Behr, P. J., Hersh, D. and Foreman, K. 1997. Macroalgal blooms in shallow estuaries: Controls and ecophysiological and ecosystem consequences. Limnology and Oceanography. doi:10.4319/lo.1997.42.5_part_2.1105

Viaroli, P., Bartoli, M., Giordani, G., Naldi, M., Orfanidis, S. and Zaldivar, J. 2008. Community shifts, alternative stable states, biogeochemical control and feedbacks in eutrophic coastal lagoons: a brief overview. Aquatic Conservation: Marine and Freshwater Ecosystems. 18: 105-117.

WFD-UKTAG (Water Framework Directive – United Kingdom Technical Advisory Group). 2014. UKTAG Transitional and Coastal Water Assessment Method Macroalgae Opportunistic Macroalgal Blooming Tool. Retrieved from http://www.wfduk.org/sites/default/files/Media/Characterisation of the water environment/Biological Method Statements/TraC Macroalgae OMBT UKTAG Method Statement.PDF.

Macrofauna

Due to their relatively sedentary life style and sensitivity to changes in pressures, soft sediment macrofauna can indicate and integrate complex environmental conditions and therefore considered useful for representing benthic community health in response to contaminants, nutrients, organic enrichment, deposition rates, turbidity and changes in muddiness (if representative sites are surveyed). Considering soft bottom macrofauna is particularly useful given the fact that most NZ estuaries are dominated by soft sediments. As an estuary progresses along the gradient of increasing eutrophication and muddiness, the benthic macroinvertebrate community responds to lowering oxygen and increasing toxicity by shifting towards smaller, more stress tolerant species. These are not as efficient at bioturbation, which limits oxygen penetration into the sediments and effectively minimise the zone of coupled nitrification/denitrification in the sediments (Pearson and Rosenberg 1978, Sutula 2011). They are also often less efficient in providing other ecosystem services, e.g., secondary production, biofiltration or provisioning. However, developing macrofauna-based attributes might be complicated by the high variability of natural conditions in estuaries and multivariate response of the macrofauna communities.

Sampling design

National standards / guidelines and consistency: No national standards. Although sampling designs for collecting fine scale macrofauna information vary across the country (i.e., site extent, number of replicates, frequency and sampling time), many current monitoring programmes rely on guidelines described in Thrush et al. (1988) or the closely related Robertson et al. (2002) protocols.

Potential bottlenecks: Selection of non-representative sampling sites, not reflecting the entire gradient of degradation/response to a stressor.

Opportunities: To optimize sampling strategy, macrofauna sampling can be aligned with collecting data on other benthic attributes and/or state variables (e.g., sediment quality characteristics, sediment and shellfish contaminants)

Caveats and recommendations: Soft bottom macrofaunal communities are considered rather stable (Turner et al. 1995), therefore there often no need for extensive temporal replication to detect significant changes in response to diffuse impacts and numerous point sources (Hewitt and Thrush 2007). However, consistent sampling times with respect to season are recommended. While winter is probably the season that would be most suitable, due to lack of recruitment over this time period, present sampling either samples throughout the year or in summer (February) or spring (October). The few programmes that sample two -to four monthly throughout the year offer the potential to reconcile these differences and to account for short-term variations. Appropriate spatial replication is recommended to account for spatial variability. Number of sampling sites per estuary does not need to be uniform, but at least several sites that are representative of habitats both highly susceptible and less susceptible to the relevant pressures should be sampled (Robertson et al. 2016a). Spatial replicates within each location should be positioned at least 5 m apart from each other to limit the influence of spatial autocorrelation (Greenfield at al. 2013). Most of the current macrofauna sampling campaigns target intertidal habitat. Although it is usually a dominant and relatively vulnerable habitat, it may not represent the whole estuary condition in response to upstream effects. Therefore, sampling subtidal sites should be considered for certain estuaries.

Sampling procedures
National standards / guidelines and consistency: No national standards. Macrofauna sampling is reasonably consistent across the country, most are following Robertson et al. (2002) guidelines – using sediment core (130 or 150 mm diameter; 100 or 150mm depth) for infauna sampling and 0.25 m^2 quadrats for epifauna. Sample preservation methods are quite variable though (different concentrations of formalin, ethanol, isopropyl alcohol, etc.) **Potential bottlenecks:** Poor preservation of samples may impede correct identification of taxa. **Opportunities:** 95-96% ethanol preserved samples can be used for molecular identification of the unknown or ambiguous specimens and ID validation of the cryptic species and juveniles. **Caveats and recommendations**: A standardized core size and depth and proper concentration of preservative used in the field would improve the quality of samples and comparability of results. The effect of differences in replication is unlikely to be an issue, however, analysis of this factor is required to be sure.
Laboratory analyses
National standards / guidelines and consistency: No national standards. Sample sieving and specimens picking approaches are rather consistent; the organisms are sorted under microscope, identified to the lowest taxonomic level possible, enumerated. However "lowest taxonomic level" can vary significantly among the labs and taxonomists. Often, larger fauna are identified internally, while small and cryptic animals sent to external taxonomic experts. Many taxa are identified to relatively broad levels of taxonomic resolution (Family, Class, Order or even Phylum), however, this is consistent with international practice. Since 2014, some monitoring protocols are following QA procedures developed by Hewitt et al. (2014) for regional councils. **Potential bottlenecks:** Limited, variable and inconsistent taxonomic resolution. Reduced abundances collected on a larger sieve size can bias the diversity assessment and mask stressor-driven changes in macrofauna communities. Damage of organisms during the sieving can affect the identification success and precision. **Opportunities:** To allow temporal-spatial comparisons, taxonomic resolution and naming can be aligned with the previous works. Rapidly evolving molecular techniques (e.g., barcoding and metabarcoding) might be employed to complement and (eventually) substitute taxonomic identification process at least for some taxa. This would allow for better resolved and standardized taxonomic IDs. **Caveats and recommendations**: Robust QA procedures are advisable at every stage of sample processing (picking, IDs, enumeration and data entries), following Hewitt et al. (2014) protocol. A well-established reference taxonomic collection is needed, and eventually - a well-annotated and continuously curated national molecular reference database, compatible with the international databases (e.g., GenBank, BOLD, PR2, etc.). Gentle sieving is recommended for reducing identification bias and if sorted individuals are preserved in 95-96% ethanol post-hoc validation (including molecular ID) can be performed when needed.
Computational approaches and metrics derived
National standards / guidelines and consistency: No national standards (might be developed for MCI soon). Various macrofaunal biodiversity metrics are being calculated in different monitoring programmes and other national initiatives; e.g., BHM (Anderson et al. 2006), ES (Keeley et al. 2012), TBI (Rodil et al. 2013), RI_AMBI (Robertson et al. 2016b). The choice of a metric is driven by the aim of a study and data type derived by particular sampling and sample processing approach (i.e., abundance, relative abundance or presence-absence). The computational approaches might also vary for the same

metric, e.g., different local adjustments applied to the indices developed overseas, tolerance rankings, etc.

Potential bottlenecks: Some metrics, especially the diversity/richness metrics and those involving tolerance values are highly dependent on taxonomic resolution and inconsistencies in the resolution across the samples (Clapcott et al. 2017).

Opportunities: Multiple biodiversity metrics (indices) can be derived from the same macrofauna sample and can be utilized for establishing responses to different stressors and setting relevant bands and thresholds. Semi-quantitative metrics can be developed based on molecular (metabarcoding) data (Aylagas et al. 2014).

Caveats and recommendations: Multivariate indices have been demonstrated to outperform simple metrics for measuring stressor gradients in NZ estuaries (Ellis et al. 2015). However, in order to develop any of the existing macrofauna metrics into robust attributes and/or state variables, comprehensive and consistent testing is needed to better understand their natural broad-scale spatial variability (i.e., bioregion, estuary, estuary type, etc.) and responses to stressors.

Thresholds

For eutrophication-related macrofauna response, thresholds have been recommended for the RI_AMBI for the shallow lagoon type estuaries (ETI Tool 2; Robertson et al. 2016a) and on national scale (Robertson et al. 2016b); and also within the Auckland region for BHM (Anderson et al. 2006). However threshold values may need to be calibrated for different stressors/ specific estuary/ estuary type/ bioregion to ensure that differences in natural variability are accounted for. Deriving standardized thresholds is impeded by the high natural variability of macrofauna communities in estuaries, limited data on reference conditions and stressor-specific response on a national scale (Berthelsen et al. in press).

Emerging and prospective future methods

Rapidly advancing molecular techniques provide promising tools for species identification and overcoming the current methodological bottlenecks related to insufficient taxonomic resolution, inconsistent IDs, overall declining morphological expertise. For example, *DNA barcoding,* based on PCR amplification and sequencing of a taxa-specific DNA fragment and assignment of taxonomy by comparing the sequence to reference databases, is applicable for verifying identity of cryptic, ambiguous or damaged specimens. Metabarcoding, combined with high-throughput sequencing, is used for characterising the biodiversity of biological communities in environmental samples. Metabarcoding assays can be designed to encompass broad taxonomic groups (e.g., eukaria) or target specific taxa, such as polychaetes. Application of metabarcoding in environmental monitoring can enable development of molecular biotic indices with higher sensitivity and specificity to anthropogenic pressures (Keeley et al. 2017). This can be achieved both by improved taxonomic resolution in molecular datasets and incorporation of taxonomic groups not commonly considered in benthic surveys, e.g., meiofauna, protists, bacteria (Pochon et al. 2016, Laroche et al. 2017). However, estimation of abundance rather than presence/absence is not yet achievable by these techniques, not are most species bar-coded.

References

Anderson, M.J., Hewitt, J.E., Ford, R.B., Thrush, S.F. (2006) Regional models of benthic ecosystem health: predicting pollution gradients from biological data. *Auckland Regional Council Technical Publication*. http://www.arc.govt.nz/shadomx/apps/fms/fmsdownload.cfm?file_uuid=03EAFF07-BCD4-1A24-933D-9092E8443D00&siteName=arc

Aylagas E, Borja A, Rodriguez-Ezpeleta N 2014. Environmental status assessment using DNA metabarcoding: towards a genetics based marine biotic index (gAMBI). PLOS One 9: e90529.

Berthelsen A, Atalah J, Clark D, Goodwin E, Sinner J 2017. Relationships between biotic indices, multiple stressor gradients and natural variability in New Zealand estuaries. Ecological Indicators (in press).

Clapcott J, Wagenhoff A, Neale M, Death R, Storey R, Smith B, Harding J, Matthaei C, Collier K, Quinn J, Young R 2017. Macroinvertebrate metrics for the National Policy Statement for Freshwater Management project: Report 1. Prepared for Ministry for the Environment. Cawthron Report No. 3012. 99 p. plus appendices.

Ellis J, Ysebaert T, Hume T, Norkko A, Bult T, Herman P, Thrush S, Oldman J 2006. Predicting macrofaunal species distributions in estuarine gradients using logistic regression and classification systems. Marine Ecology Progress Series 316: 69-83.

Greenfield B, Hewitt J and Hailes S 2013. Manukau Harbour ecological monitoring programme: report on data collected up until February 2013. Prepared by NIWA for Auckland Council. Auckland Council technical report, TR2013/027. 51p.

Hewitt JE, Hailes SF, Greenfield BL 2014. Protocol for Processing, Identification and Quality Assurance of New Zealand Marine Benthic Invertebrate Samples. NIWA Report NRC15203. 36p.

Hewitt JE, Thrush SF 2007. Effective long-term ecological monitoring using spatially and temporally nested sampling. Environ Monit Assess 133: 295-307.

Keeley N, Forrest B, Crawford C, Macleod C 2012. Exploiting salmon farm benthic enrichment gradients to evaluate the regional performance of biotic indices and environmental indicators. Ecological Indicators 23: 453-466.

Keeley N, Wood SA, Pochon X 2017. Development and validation of a multi-trophic metabarcoding biotic index for monitoring benthic organic enrichment. Ecological Indicators (in press).

Laroche O, Wood SA, Tremblay LA, Ellis JI, Lear G, Pochon X 2017. A cross-taxa study using environmental DNA/RNA metabarcoding to measure biological impacts of offshore oil and gas drilling and production operations. Marine Pollution Bulletin (in revision).

Pochon X, Cornelisen C, Morrisey D, Wood S 2016. Survey of microbial seabed communities in the Firth of Thames and Hauraki Gulf using a DNA metabarcoding approach. Prepared for Waikato Regional Council. Cawthron Report No. 2880. 22 p., plus appendix.

Pearson TH, Rosenberg R 1978. Macrobenthic succession in relation to organic enrichment and pollution in the marine environment. Oceanography and Marine Biology: an Annual Review 16: 229-311.

Robertson BM, Gillespie PA, Asher RA, Frisk S, Keeley NB, Hopkins GA, Thompson SJ, Tuckey BJ 2002. Estuarine Environmental Assessment and Monitoring: A National Protocol. Part A. Development, Part B. Appendices, and Part C. Application. Prepared for supporting Councils and the Ministry for the Environment, Sustainable Management Fund Contract No. 5096. Part A. 93p. Part B. 159p. Part C. 40p. plus field sheets.

Robertson BM, Stevens L, Robertson B, Zeldis J, Green M, Madarasz-Smith A, Plew D, Storey R, Oliver M 2016a. NZ Estuary Trophic Index Screening Tool 2. Determining Monitoring Indicators and Assessing Estuary Trophic State. Prepared for Envirolink Tools Project: Estuarine Trophic Index, MBIE/NIWA Contract No: C01X1420. 68p.

Robertson BP, Savage C, Gardner JPA, Robertson BM, Stevens LM 2016b. Optimising a widely-used coastal health index through quantitative ecological group classifications and associated thresholds. Ecological Indicators 69: 595-605.

Rodil IF, Lohrer AM, Hewitt JE, Townsend M, Thrush SF, Carbines M 2013. Tracking environmental stress gradients using three biotic integrity indices: Advantages of a locally-developed traits-based approach. Ecological Indicators 34, 560-570.

Sutula M 2011. Review of indicators for development of nutrient numeric endpoints in California estuaries. Prepared for: The California Environmental Protection Agency State Water Resources Control Board (Agreement Number 07-110-250), 07.

Thrush, S. F., Pridmore, R. D., Hewitt, J. E., Roper, D. S. 1988. Design of an Ecological Monitoring Programme for the Manukau Harbour. Report prepared by the Water Quality Centre for Auckland Regional Water Board, Auckland Regional Authority. Technical Publication Number 79: 62.

Turner SJ, Thrush SF, Pridmore RD, Hewitt JE, Cummings VJ, Maskery M 1995. Are soft-sediment communities stable? An example from a windy harbour. Marine Ecology Progress Series 120: 219-230.

Sediment Chlorophyll a

Nutrients flushed from the land and transported down streams/rivers into estuaries can be utilised by benthic microalgae growing on sediments in estuarine intertidal flats. These microalgae often take advantage of excess nutrients. In extreme cases of nutrient loading, dense green to orange films or mats can be seen covering the sediment surface. Chlorophyll-a is a primary photosynthetic pigment contained in microalgae (as well as other plants) (Robertson et al. 2002). The amount of Chlorophyll-a within sediments can be an indicator of microalgal mat cover and, more specifically, photosynthetically active microalgal biomass (Robertson et al. 2002). However, sediment chlorophyll-a can decrease due to reduced light availability / sediment resuspension with sediment loading. Sediment chlorophyll-a is relatively easy to measure. Sediment chlorophyll-a content can also provide information regarding the condition or state of estuaries and link to values such as ecosystem health. Although microalgae provide food for benthic animals including many shellfish, microalgal blooms and/or mats can be indicative of eutrophic (highly enriched) conditions (Robertson et al. 2002). There are some issues to consider if using sediment chlorophyll-a as an attribute for managing upstream effects. Changes in nutrients, and therefore potentially sediment chlorophyll-a, can result from other sources besides rivers/streams e.g., storm and wastewater outfalls, oceanic inputs. Natural variability may also confound changes in state associated with upstream effects, and microalgal densities are known to be inherently extremely variable (Gillespie et al. 2009).
Sampling design
<u>National standards / guidelines and consistency:</u> No national standards. Although sampling designs for collecting sediment chlorophyll-a data vary across the country (i.e., site extent, number of replicates, frequency and sampling time, site representativeness), some council-led intertidal monitoring programmes follow EMP (Robertson et al. 2002) guidelines. Sampling of sediment chlorophyll-a is largely conducted in intertidal habitats although, in some regions, subtidal sites are also included. **<u>Potential bottlenecks:</u>** Inappropriate sampling design, in terms of spatial extent and number of replicates, selection of non-representative sampling sites, not reflecting the entire gradient of degradation/response to a stressor. Differences in sampling frequency and timing. Possible differences in representativeness in regards to location of sample collection e.g., " as microalgal densities are known to be inherently extremely variable, core positions were intentionally selected to sample regions of visible yellow/green coloration in order to estimate maximum chlorophyll-a concentrations" (Gillespie et al. 2009) as opposed to "on each sampling occasion surface scrapes were collected and analysed for chlorophyll-a with permanent monitoring plots randomly located at the mid-intertidal level at each site" (Singleton 2010). **<u>Opportunities:</u>** To optimize sampling strategy, sediment chlorophyll-a sampling can be aligned with collecting data on other benthic attributes and/or state variables (e.g., sediment quality characteristics, macrofauna, sediment and shellfish contaminants). **<u>Caveats and recommendations</u>**: Appropriate spatial replication is recommended to account for spatial variability. Most of the current programmes target intertidal habitat. Although it is usually a dominant and relatively vulnerable habitat, it may not represent the whole estuary condition in response to upstream effects. Therefore, sampling subtidal sites should be considered.

Sampling procedures
National standards / guidelines and consistency: No national standards. Sampling procedures appear generally similar, with samples collected and then frozen prior to analysis. **Potential bottlenecks:** There may be differences in the sample size (e.g., including surface scrape, core, or syringe), sample depth (e.g., including surface, 2 cm or 5 cm), and the treatment of samples (e.g., including compositing and sub-sampling). Differences also exist in the storage of samples (e.g., some specify that samples are kept in dark container or kept on ice prior to freezing) and the timing of analysis e.g., some specify that analysis must occur within 1 month of collection, while others don't specify a timeframe. **Caveats and recommendations**: Standardise sampling procedures in regards to bottlenecks.
Laboratory analyses
National standards and consistency: No national standards. Different analyses have been used, with examples including the NIWA Periphyton Monitoring Manual and Limnology and Oceanography 1967 No 12. Results from these are not necessarily comparable. **Potential bottlenecks:** Analysis results are not necessarily comparable. **Opportunities:** Data from the analysis of phaeophytin (plant degradation products) can also be collected (Robertson et al. 2002), and often is. The EMP suggests to preserve additional samples with Lugol's iodine solution for later microscopic identification dominant taxa (Robertson et al. 2002).
Computational approaches and metrics derived
National standards / guidelines and consistency: No national standards. **Potential bottlenecks:** Chlorophyll-a results are often reported in different units. **Caveats and recommendations**: Standardise sediment chlorophyll-a units for direct comparability of results.
Thresholds
None. Any thresholds developed may need to be calibrated for different stressors/ specific estuary/ estuary type/ bioregion to ensure that differences in natural variability are accounted for. Deriving standardized thresholds may be impeded by natural temporal variability.
References
Gillespie P, Clement D, Asher R 2009. Delaware Inlet fine scale benthic baseline 2009. Prepared for Nelson City Council. Cawthron Report No. 1594. 28.p. Robertson BM, Gillespie PA, Asher RA, Frisk S, Keeley NB, Hopkins GA, Thompson SJ, Tuckey BJ 2002. Estuarine Environmental Assessment and Monitoring: A National Protocol. Part A. Development, Part B. Appendices, and Part C. Application. Prepared for supporting Councils and the Ministry for the Environment, Sustainable Management Fund Contract No. 5096. Part A. 93p. Part B. 159p. Part C. 40p. Singleton N 2010. Regional estuary monitoring programme (REMP) data report: benthic macrofauna communities and sediments – July 2007 – April 2008. Southern Firth of Thames and Whaingaroa (Raglan) Harbour. Prepared for Environment Waikato. Document #: 1520788.

Mud content/ grain size

Mud/sand habitat is often the dominant habitat type in New Zealand estuaries (Robertson et al. 2002). Changes in sediment grainsize can be indicative of habitat change and type of sediment supply, and can occur as a result of terrestrial sediment (Hewitt et al. 2014). Although natural to some extent, the amount of land-derived fine sediments entering estuaries has increased due to anthropogenic impacts associated with changes in land use (e.g., deforestation). Through land run-off, fine sediments enter streams and rivers and eventually be deposited into estuaries if not flushed out into coastal waters (Robertson et al. 2002). Sediment mud content within estuaries can increase as a result of fine sediments entering estuaries, and hence can be used as a surrogate for sediment accumulation (Hewitt et al. 2014). Townsend and Lohrer (in prep.) state that bed-sediment mud content is relatively simple and cheap to measure, and is already a component of many monitoring programmes. Sediment mud content can also provide information regarding the condition or state of estuaries and link to values such as ecosystem health. For example, in New Zealand sediment mud content it is a major stressor of benthic animal communities, such as macrofauna (Robertson et al. 2015) including mahinga kai species, and ecological responses to bed-sediment mud content are reasonably well understood (Gibbs and Hewitt 2004, Thrush et al 2004). Underfoot condition (muddiness) is also a key component in human preference and the value people place on marine environments (Batstone & Sinner 2010). Fine sediments can become contaminated with elevated nutrients, organic matter, potentially disease-causing organisms and potentially toxic chemicals from anthropogenic activities (Robertson et al.. 2002). The tendency for sediments to become anoxic is also higher if the sediments are muddy and interstitial spaces small (ETI 2; Robertson et al. 2016a). There are some issues to consider if using mud content as an attribute for managing upstream effects. Besides terrestrial sediment inputs from rivers/streams, changes in sediment grain size can result from human activities such as mining, bottom fishing and dumping of dredge disposal (Hewitt et al. 2014). Sediment mud content can also be influenced by natural processes, such as resuspension, remobilization by currents and water movement, occurring within estuaries and associated coastal waters (Hewitt et al. 2003). Influence of these *in-situ* impacts/ processes on sediment mud content may confound upstream effects. Natural variability may also confound changes in state associated with upstream effects e.g., a number of processes that can cause high natural variability in sedimentation have been identified (Townsend & Lohrer, in prep.), and natural within-year and between-year variability in sediment grainsize without a strong predictable pattern has been documented for some intertidal areas (Hewitt et al. 2014).
Sampling design
National standards / guidelines and consistency: No national standards. Although sampling designs for collecting mud content data vary across the country (i.e., site extent, number of replicates, frequency and sampling time, site representativeness), many current monitoring programmes in intertidal habitats rely on guidelines described in Robertson et al. (2002). Although most council-led estuary monitoring programmes focus on intertidal habitats, subtidal sediment samples are also collected in New Zealand estuaries for a variety of reasons (e.g., SOE, consent monitoring). Sampling design for subtidal sample collection varies between programmes. **Potential bottlenecks:** Inappropriate sampling design, in terms of spatial extent and number of replicates, selection of non-representative sampling sites, not reflecting the entire gradient of degradation/response to a stressor. Differences in sampling frequency – note that without some continuously monitored sites, temporal cycles related to ENSO may result in detection of changes that are the result of natural cycles. In the case of a national monitoring strategy, monitored sites need to

be spatially distributed around the country and at present continuous monitoring of some sites only occurs in the northern half of the North Island (Hewitt et al. 2014). **Opportunities:** To optimize sampling strategy, mud content sampling can be aligned with collecting data on other benthic attributes and/or state variables (e.g., macrofauna, sediment and shellfish contaminants). **Caveats and recommendations**: Analyses present data to understand the effect of temporal frequency and spatial replication. Assess the value of reporting mud content at current sampling sites as an indicator of the ecological issue of sedimentation, noting that monitoring sites may be biased towards high risk areas of estuaries (Bolton-Richie & Lawton in draft). Most of the current sediment grain size sampling programmes target intertidal habitat. Although it is usually a dominant and relatively vulnerable habitat, it may not represent the whole estuary condition in response to upstream effects. Therefore, sampling subtidal sites should be considered for certain estuaries.
Sampling procedures
National standards / guidelines and consistency: No national standards. Mud content sampling in intertidal habitats is reasonably consistent across the country, most are following Robertson et al. (2002) guidelines – using 0-20 mm sediment sample depth (Bolton-Richie & Lawton in draft). The number (if any) of samples composited prior to analyses also varies across monitoring programs. Subtidal sampling is often conducted using divers to collect cores or from a vessel using a grab. **Potential bottlenecks:** Differences in sample depth and number of samples composited. **Opportunities:** Paired (double) samples could be collected, with one sample from each pair analysed using a different laboratory analysis in order to obtain data on the comparability of results from these analyses (see Laboratory analyses section below). **Caveats and recommendations**: Standardize sample depth and, possibly, sample compositing.
Laboratory analyses
National standards / guidelines and consistency: No national standards. Analysis methods for sediment particle size distribution fall into two main categories: Laser Diffraction Particle Analysis and Wet Sieving, gravimetric (calculation by difference) with some variation in analysis methods occurring within these two methods (Bolton-Richie & Lawton in draft). **Potential bottlenecks:** Results between and within two main grain size analysis methods are not necessarily comparable (Bolton-Richie & Lawton in draft). For example, grain size results obtained through sieving can be fuzzy as sieve mesh pores are square, resulting in different sized particles passing through depending on whether they are spherical or elongated (Hewitt et al. 2014). Also, laser diffraction particle analysis does not measure full range of grain sizes, e.g., larger (>2000um) grain sizes are excluded and in some cases <3.9 um as well. As mud content is analysed as a proportion of total sediment, it is not always clear what the 'total' represents in regards to grain size. **Caveats and recommendations**: Because none of the methods has a clear advantage over the others, research is needed to determine if measured degrees and rates of change are similar across methods (Hewitt et al. 2014). Ultimately, to allow future temporal-spatial comparisons, a standardized grain size analysis method should be decided on and used throughout New Zealand.

Computational approaches and metrics derived
<u>National standards / guidelines and consistency:</u> No national standards. Most intertidal (and subtidal) monitoring programmes report as in Robertson et al. (2002), where mud is defined as grain sizes <62.5 µm and expressed as a proportion of total sediments. The ETI is also consistent with this (ETI-2; Robertson et al. 2016a). **<u>Potential bottlenecks:</u>** Mud is often referred to as a combination of smaller sediment fractions silt and clay. However, during reporting it is not always made clear what grain size fractions these three terms represent, making it difficult to compare results. Depending on the definition of mud in regards to grain size and the laboratory analysis used, results from smaller grain size fractions (e.g., often referred to as silt and clay) may need to be combined to produce an overall value that represents mud. **<u>Caveats and recommendations</u>**: Standardize the definition of mud in regards to grain size (e.g., as in Townsend & Lohrer, in prep.). This will help to ensure comparability in reporting of sediment mud content.
Thresholds
Estuarine Trophic Index (ETI) assigns risk rating thresholds to % mud content for individual sites in regards to eutrophication status (ETI-2; Robertson et al. 2016a). Interim thresholds, pending further research, are currently proposed within this index for New Zealand shallow lagoon type estuaries. Auckland council have also developed a muddiness scale against which sediment mud content data can be compared (cited in Bolton-Richie & Lawton in draft). Thresholds may need to be calibrated for different stressors/ specific estuary/ estuary type/ bioregion to ensure that differences in natural variability are accounted for. For example, Townsend and Lohrer (in prep.) describe estuary typology classification for sedimentation. Deriving standardized thresholds may be impeded by natural temporal variability (within-year and between year) without a strong predictable pattern in sediment grain size that has been documented for some intertidal areas (Hewitt et al. 2014). Identification of reference conditions could assist in the derivation of ecological health thresholds for different regions and estuary types. A guideline could involve maintaining bed-sediment mud content below a critical value across a specified areal proportion of an estuary. There is a good understanding of how benthic communities and functional health change over mud gradients (Hewitt et al. 2012) and this could be used to derive the guideline. For example, Rodil et al. (2013) demonstrate how functional health changes in relation to mud content (index scores were highest below 10% mud and always low above 60% mud). Nevertheless, converting this information into guidelines for areal extents in estuaries would require a process similar to the one undertaken for sedimentation (e.g., literature review, data mining, workshopping, consultation, peer review). This threshold urgently requires further development but is proposed as a key indicator as changes to sediment mud content, a known driver of ecological shifts, can occur without being detected by other indicators (see 2-4 in Sediment Deposition Rate Methods Factsheet). Initial bottom line guidance could be 'sediment mud content at representative sites should not increase from its current extent'.
References
Batstone, C., Sinner, J. (2010) Techniques for evaluating community preferences for managing coastal ecosystems. Auckland region stormwater case study, discrete choice model estimation. Prepared by Cawthron Institute for Auckland Regional Council. *Auckland Regional Council Technical Report 2010/012.* Bolton-Ritchie L, Lawton R (in draft). Summary of Regional Council metadata collected to inform the coastal ecosystem module of LAWA. Gibbs, M., Hewitt, J.E. (2004) Effects of sedimentation on macrofaunal communities: a synthesis of research studies for ARC. Auckland Regional Council, Auckland.

Hewitt JE, Lohrer AM, Townsend M 2012. Health of estuarine soft-sediment habitats: continued testing and refinement of state of the environment indicators. Prepared by NIWA for Auckland Council. *Auckland Council technical report*, TR2012/012

Hewitt, J.E., Bell, R., Costello, M., Cummings, V., Currie, K., Ellis, J., Francis, M., Froude, V., Gorman, R., Hall, J., Inglis, G., MacDiarmid, A., Mills, G., Pinkerton, M., Schiel, D., Swales, A., Law, C., McBride, G., Nodder, S., Rowden, A., Smith, M., Thompson, D., Torres, L., Tuck, I., Wing, S. (2014) Development of a National Marine Environment Monitoring Programme (MEMP) for New Zealand.

Hewitt, J.E., Cummings, V.J., Ellis, J.I., Funnell, G.A., Norkko, A., Talley, T.S., Thrush, S.F. (2003) The role of waves in the colonisation of terrestrial sediments deposited in the marine environment. Journal of Experimental Marine Biology and Ecology, 290: 19-47.

Robertson BM, Gillespie PA, Asher RA, Frisk S, Keeley NB, Hopkins GA, Thompson SJ, Tuckey BJ 2002. Estuarine Environmental Assessment and Monitoring: A National Protocol. Part A. Development, Part B. Appendices, and Part C. Application. Prepared for supporting Councils and the Ministry for the Environment, Sustainable Management Fund Contract No. 5096. Part A. 93p. Part B. 159p. Part C. 40p. plus field sheets.

Robertson, B.P., Gardner, J.P.A., Savage, C., 2015. Macrobenthic-mud relations strengthen the foundation for benthic index development: A case study from shallow, temperate New Zealand estuaries. Ecological Indicators 58, 161-174.

Robertson BM, Stevens L, Robertson B, Zeldis J, Green M, Madarasz-Smith A, Plew D, Storey R, Oliver M 2016a. NZ Estuary Trophic Index Screening Tool 2. Determining Monitoring Indicators and Assessing Estuary Trophic State. Prepared for Envirolink Tools Project: Estuarine Trophic Index, MBIE/NIWA Contract No: C01X1420. 68p.

Rodil, I.F., Lohrer, A.M., Hewitt, J.E., Townsend, M., Thrush, S.F., Carbines, M. (2013) Tracking environmental stress gradients using three biotic integrity indices: Advantages of a locally-developed traits-based approach. *Ecological Indicators*, 34: 560–570.

Townsend M, Lohrer AM (in prep.). ANZECC Guidance for Estuary Sedimentation. NIWA CLIENT REPORT No: HAM2015-096. Prepared for Ministry for the Environment.

Thrush, S.F., Hewitt, J.E., Cummings, V.J., Ellis, J.I., Hatton, C., Lohrer, A., Norkko, A. (2004) Muddy waters: elevating sediment input to coastal and estuarine habitats. Frontiers in Ecology and Environment, 2: 299-306.

Deposition rate

Sediment

Suspended sediment and fine sediment deposition (e.g., particles <0.0625 mm in diameter) are recognised as significant threats to estuaries and coastal environments in many parts of the world (e.g., McKnight 1969, Woods and Armitage 1997, Thrush et al. 2004). In some estuaries, particularly shallow intertidally dominated ones, land disturbance in the catchment can result in increased fine sediment mobilisation, resulting in significant mud deposition zones in the upper estuary tidal flats (Robertson et al. 2016b). Townsend and Lohrer (in prep.) report consequences from short-term "event" sedimentation (primarily burial) that include lethal effects on benthic biota, changes in benthic species composition, loss of sensitive species, decline in diversity, and modification of animal behaviours (Hewitt et al. 2003, Thrush et al. 2004, Lohrer et al. 2004, Norkko et al. 2002a). It can also alter microbial activities (which are critical for organic matter degradation and nutrient regeneration), diminish benthic primary productivity, and reduce the oxygenation of surficial sediments (by capping the seabed, clogging sediment pore spaces, and depriving micro- and macrophytes of light) (Berkenbusch et al. 2002). An added consideration is that fine sediment loads are often accompanied by elevated nutrient loads. Their combined effect can cause sediments to become eutrophic (Robertson and Stevens 2012, 2013, Robertson et al. 2016b). The resulting "soft mud/macroalgae cocktail" exacerbates sediment deoxygenation, production of sulphides, and degraded macrobenthos.

The consequences of longer term sedimentation on estuarine communities (over months or years) are not as well studied (Anderson et al. 2004, Townsend et al. 2014), but high rates of sedimentation are capable of altering estuarine habitats, modifying ecosystem functions and decreasing a broad range of ecosystem services. Extensive NEMP monitoring data from typical NZ shallow tidal lagoon, tidal river and ICOLL estuaries show that extensive areas of soft mud, elevated sedimentation rates, and high sediment mud contents are commonly associated with a degraded macroinvertebrate community, and particularly so where nutrients are excessive and soft mud areas are overlain with dense nuisance beds of opportunistic macroalgae (Robertson et al. 2016b). Further, legacy effects of previous land use decisions mean sediment impacts may be borne for decades or centuries after management changes are made, with some effects nearly impossible to reverse (Townsend and Lohrer, in prep.).

For these reasons, mud is considered a key attribute for management and a useful supporting indicator for the assessment of estuary trophic status (i.e., if soft muds are present then the estuary is more prone to eutrophic sediments).

Townsend and Lohrer (in prep.), in their MFE review report providing ANZECC guidance for estuary sedimentation, concluded that a standalone measure of annual sedimentation rate would be insufficient for managing sediment effects in estuaries. However, it may provide benefit as a foundation for a broader framework that includes other elements related to sediment stress, such as suspended sediment concentration (SSC), bed sediment particle size distribution (for mud content), and the areal extent of muddy sediment in an estuary. Measures of these elements, supported by extensive NZ estuary monitoring data, were included in the ETI (Robertson et al. 2016b).

Based on the above work, a multi-faceted approach is recommended that includes mud content, mud sedimentation rate, and the spatial distribution of these be used to assess sediment condition (and the trophic state) of shallow (<3 m mean depth), tidal lagoon, tidal river and ICOLL type estuaries. Such indicators will monitor the infilling rate, whether there has been a shift to finer sediments, and the spatial extent of any changes. Supporting state variables should include monitoring of plants and animals so that the effects of mud changes on key biota (e.g., macroinvertebrates, fish, seagrass) can be gauged, as well as ensuring water clarity is not adversely impacted by suspended fine sediments.

The following specific measures of estuary sediment are proposed, based on work undertaken in development of ANZECC guidance for estuary sedimentation (Townsend and Lohrer, in prep.) which focused on the physical impact of sediment accumulation on the benthos, development of the ETI (Robertson et al. 2016b) which addressed sediment related impacts associated with eutrophication impacts, and MPI (2016) which provided an analysis of mitigations to manage sediment and *E. coli* loads in the Whangarei Harbour catchment in order to meet freshwater objectives and limits.

Annual average sedimentation rate (AASR)

MPI (2016) proposed use of an annual-average sedimentation rate (AASR). This is defined as: *Mass of sediment deposited per year/(settled-sediment density multiplied by the area over which sediment deposits).* Using a simple parameter such as the AASR means it is relatively easy to measure and explain progress towards achieving it. It may also be a suitable "master attribute" that is indicative of a wide range of sediment effects in estuaries. The AASR is unambiguous, readily measurable (by, for example, repeat bathymetric surveys or sedimentation plates) and easy to relate to catchment sediment inputs (Green, 2013).

Townsend and Lohrer (in prep.) provide some qualifiers to the use of a single metric highlighting that "average sedimentation" for the whole estuary is difficult to interpret meaningfully. An estuary with an "overall" average sedimentation rate below a set guideline may still contain multiple sites where the levels are exceeded, while the inclusion of estuary areas with low sedimentation will reduce and 'dilute' the magnitude of the overall sedimentation rate, potentially obscuring instigation of necessary management responses. They recommended examining estuarine sites individually, or by category, to then initiate a proportionate management response following a review of the data.

To protect against significant adverse impacts from future event-scale effects, Townsend and Lohrer (in prep.) proposed a default guideline value for sedimentation of no more than 2mm of sediment accumulation per year above the natural annual sedimentation rate for the estuary, or part of estuary, at hand. Such a measure is expected to provide reasonable protection to sediment macrofauna in deposition zones from physical impacts (primarily burial), however Townsend and Lohrer (in prep.) emphasise it does not take into account indefinite resilience (which refers to the ability of an environment to absorb a given amount of a stressor in perpetuity, rather than in a time-bound capacity - Kelly et al. 2015). Nor does it take into account the natural sedimentation rate of the estuary, or the extent of change from natural state the estuary may have already undergone. To address these limitations it is proposed that changes from natural state are also incorporated as follows:

Annual average sedimentation ratio (AAS Ratio)

The ETI proposed a simple metric to manage sediment inputs based on the natural sedimentation rate (NSR) of the estuary. This is because estuaries with different catchment geologies and erosion rates have a different natural sensitivity to sediment inputs, and consequently a universal AASR of 2mm may not reflect an appropriate management threshold in all estuaries. The NSR is the sedimentation rate for the estuary in its natural state (i.e., pre-human vegetation cover and wetland presence). This rate can be estimated as *the current sedimentation rate (CSR) multiplied by the natural state sediment load (NSL)/current sediment load (CSL) ratio*. Catchment models (e.g., CLUES) can be used to estimate NSL and CSL. CSR can also be directly measured using sediment plates and/or bathymetric methods. A more robust approach would be to use hydrodynamic modelling methods to predict estuary retention and to replace NSL and CSL with retained NSL and retained CSL. The ETI (Robertson et al. 2016b) has proposed estuary thresholds based on the ASS Ratio e.g., a mean sedimentation rate of greater than five times the natural sedimentation rate (i.e., CSR >5 x NSR mm/yr) is rated POOR (Band D).

The proportion of the estuary area with sedimentation rates >5 x the NSR mm/yr

Because sediment deposition and retention is not uniform within an estuary, there may be multiple sites within an estuary where impacts are most concentrated. Further, because soft muds are generally associated with increased organic content, nutrients, and decreased sediment oxygenation when compared to sandier sediments, the early identification and management of excessive sediment deposition is also a critical component in managing eutrophication impacts. On the basis that it is obvious that extensive areas of excessive sediment accumulation will cause ecological damage, the ETI (Robertson et al. 2016b) used expert opinion to propose estuary thresholds based on the percentage of estuary area where CSR >5 x NSR mm/yr to highlight where there is the potential for the rapid accumulation of sediments above a rate that an estuary can readily assimilate. Because the relationship between the spatial extent of muddy sediment and overall biological impacts is still being established for NZ estuaries and requires further refinement, minimum initial bottom line guidance would be ‘the proportion of estuary area with sedimentation rates >5x the NSR mm/yr should not increase from its current extent’.

This measure requires either measurement or modelling of estuary deposition zones and rates of accumulation. Many councils have sufficient information to define these areas based on broad scale substrate maps and sedimentation rate data.

Percentage of the intertidal estuary area dominated by soft mud (sediments with >25% mud content)

Townsend and Lohrer (in prep.) describe the ecological implications of increasing mud content (sediment “muddiness”) as the loss of mud-sensitive species from benthic communities, reduced biodiversity, the loss of large functionally important species, reduced functional redundancy, and altered biogeochemical fluxes and cycles (Thrush et al. 2004 and references therein, Pratt et al. 2014, Hewitt et al. 2012).

Standard measures of spatial distribution of mud habitat have been established under the NEMP (Robertson et al. 2002). However, although there is a strong relationship between increasing sediment mud content and persistent ecological degradation (e.g., to macrofauna - Robertson et al. 2015), the relationship between the spatial extent of muddy sediment and overall biological impacts is still being established for NZ estuaries.

On the basis that it is obvious that extensive areas of soft mud will cause ecological damage, the ETI used expert opinion to conclude that if >15% of an estuary’s area is soft mud, then a high impact threshold has been breached, but noting further work was required in order to determine an overall estuary rating for soft mud. While the ETI proposes interim thresholds, minimum initial bottom line guidance would be ‘the areal coverage of muddy substrate in an estuary should not increase from its current extent’.

Sampling design
<u>National standards / guidelines and consistency</u>: There are currently no national standards for measuring sediment deposition. Broad scale spatial mapping is described in the NEMP but this requires updating to reflect subsequent advances in the protocol, and in particular the application of complimentary fine scale measures to delineate substrate boundaries and validate substrate classifications. **<u>Potential bottlenecks:</u>** Effective management requires robust information on predicted sediment yields from different land use categories and land management initiatives, and timely information on land use changes. There is currently a time lag in the release of nationally consistent data sets for land cover (e.g., LCDB), and poor consensus on which national scale models to use to estimate sediment loads (e.g., CLUES vs SedNet). Inappropriate sampling design is also a significant issue, particularly selection of non-representative sampling sites (not reflecting the entire gradient of degradation/response to a stressor), insufficient spatial coverage and replicates to characterise variation, and insufficient temporal scale measurements (sediment deposition primarily reflects pulsed rather than constant inputs so requires a long-term monitoring commitment and data record). **<u>Opportunities:</u>** Improvements to existing linked catchment and estuary models e.g., CLUES estuaries and integration of bathymetric data, hydrodynamic models, spatial maps and measured deposition rates. Increased spatial coverage and frequency of sedimentation rate measures. **<u>Caveats and recommendations</u>**: Confidence intervals on estimates are critical in the assessment of deposition rates and temporal and spatial changes. The duration of monitoring records for measured changes in bed height is also critical in assessing mean annual sedimentation rates, as is the need to relate changes to significant influences e.g., flood frequency and magnitude, within estuary redistribution, land use changes.
Sampling procedures
<u>National standards / guidelines and consistency</u>: No national standards. Measurements of changes in bed height use various techniques, materials, replicates, spatial coverage and frequencies. Although broad scale mapping generally follows the NEMP, there is inconsistency in the way data are presented and summarised making merging of data difficult. Ground truthing is also variable between different providers. As noted above, the NEMP requires updating to reflect subsequent advances in the protocol, and in particular the application of complimentary fine scale measures to delineate substrate boundaries and validate substrate classifications. Townsend and Lohrer (in prep.) provide a detailed summary of methods used to measure sediment deposition including sediment coring and dating methods (isotope tracing, caesium-137, lead-210, carbon-dating, pollen-dating), changes in bed height from a known reference point (e.g., sediment rods, traps, plates), changes in bed height across transects or over defined areas of the seabed (e.g., beach transects, bathymetric surveys, LIDAR, RTK). They conclude "there is no single measurement technique that stands out in superiority; all methods have weakness or flaws in different situations.

<u>Potential bottlenecks:</u> Establishment of bed change methods is the starting point of a future-focused long term monitoring approach and requires time to establish meaningful trends (minimum 5-10 years), and should ideally be complimented with retrospective estimates (e.g., sediment core analyses). **<u>Opportunities:</u>** Improving accuracy and decreasing cost of remote assessment of sediment bathymetry e.g., LIDAR, hydrographic surveys, which provide widespread spatial coverage on contrast to site specific measures. **<u>Caveats and recommendations</u>**: Confidence can be increased by the use of multiple complementary methods. A combination of fine-scale and broad-scale approaches would help to evaluate sedimentation over multiple spatial and temporal scales and to build a greater portfolio of information for assessing the need for management intervention (through application of the ANZECC guideline) Catchment sediment load estimates are difficult and expensive to validate. There are many unknown or poorly defined influencing factors including specific rates of sediment delivery following different types of land disturbance, sediment bed load erosion, sediment retention within estuaries, long-term cycles and influences related to climate cycles (e.g., el Niño/la mina), climate change (increased storm intensities), and human flow related changes (e.g., irrigation, flood control, dams). An important factor in determining methodology may be cost vs uncertainty, particularly if the method is likely to be the subject of Environmental Court action.
Laboratory analyses
<u>National standards / guidelines and consistency:</u> No national standards. Obtaining representative field samples is the most significant influence on analytical results. Most grain size analyses rely on standard wet sieving or laser particle analytical methods. Coring and dating methods (isotope tracing, caesium-137, lead-210, carbon-dating, pollen-dating) are all well established. **<u>Potential bottlenecks:</u>** Cost and capacity. Coring and dating analyses are relatively expensive (particularly where replicated). While grain size analyses are individually relatively cheap, spatial replication and field costs can become significant. There are limited providers of coring and dating analyses. Analytical backlogs are a potential problem given a lack of local alternative providers. **<u>Opportunities:</u>** Increased demand for coring and dating analyses may result in improved local capacity. **<u>Caveats and recommendations</u>**: Exploration of international testing options may improve capacity and reduce costs.
Computational approaches and metrics derived
<u>National standards / guidelines and consistency:</u> No national standards. The metrics are described in the first section of this factsheet. **<u>Potential bottlenecks:</u>** Sediment deposition metrics need to be related to specific estuary conditions and a sufficient monitoring interval is needed to establish robust trends. Management also requires robust estimates of sediment inputs and the ability to predict change in response to management initiatives.

<u>Opportunities:</u> Refinement of existing models to reduce uncertainty and increase accuracy of predictions. Collation of national data to enable refinement of proposed thresholds for management.
Thresholds
No national standards exist although a proposed default ANZECC Guideline Value for sedimentation of 2 mm of sediment accumulation per year above the natural annual sedimentation rate for the estuary, or part of estuary, at hand has been proposed (Townsend and Lohrer, in prep.). The ETI presents thresholds for: 1. Annual average sedimentation ratio (AAS Ratio). 2. The proportion of the estuary area with sedimentation rates >5 x the NSR mm/yr. 3. Percentage of the intertidal estuary area dominated by soft mud (sediments with >25% mud content). These thresholds require refinement and also need to be related to different estuary types.
References
Anderson, M.J., Ford, R.B., Feary, D.A., Honeywill, C. (2004) Quantitative measures of sedimentation in an estuarine system and its relationship with intertidal soft-sediment infauna. *Marine Ecology Progress Series*, 272: 33-48. Berkenbusch, K., Thrush, S., Hewitt, J.E., Ahrens, M., Gibbs, M., Cummings, V.J. (2002) Impact of thin deposits of terrigenous clay on benthic communities. Report prepared by NIWA for Auckland Regional Council. *Auckland Council Technical Publication*, 161: 56. Hewitt, J E., Lohrer, A M, Townsend, M. (2012). Health of estuarine soft-sediment habitats: continued testing and refinement of state of the environment indicators. Prepared by NIWA for Auckland Council. *Auckland Council technical report*, TR2012/012. Hewitt, J.E., Cummings, V.J., Ellis, J.I., Funnell, G., Norkko, A., Talley, T.S., Thrush, S.F. (2003) The role of waves in the colonisation of terrestrial sediments deposited in the marine environment. *Journal of Experimental Marine Biology and Ecology*, 290: 19–47. Hewitt, J E., Lohrer, A M, Townsend, M. (2012). Health of estuarine soft-sediment habitats: continued testing and refinement of state of the environment indicators. Prepared by NIWA for Auckland Council. *Auckland Council technical report*, TR2012/012. Kelly, R.P., Erickson, A.L., Mease, L.A. (2015) How not to fall off a cliff, or, using tipping points to Improve Environmental Management. *Ecology Law Quarterly*, 41(4): 843-886. Lohrer, A.M., Thrush, S.F., Hewitt, J.E., Berkenbusch, K., Ahrens, M., Cummings, V.J. (2004) Terrestrially derived sediment: response of marine macrobenthic communities to thin terrigenous Deposits. *Marine Ecology Progress Series*, 273: 121–138. Norkko, A., Thrush, S.F., Hewitt, J.E., Cummings, V.J., Norkko, J., Ellis, J.I., Funnell, G.A., Schultz, D., MacDonald, I. (2002a) Smothering of estuarine sandflats by terrigenous clay: the role of wind-wave disturbance and bioturbation in site-dependent macrofaunal recovery. *Marine Ecology Progress Series,* 234: 23-41. McKnight D.G. (1969) A recent, possible catastrophic burial in a marine molluscan community. *New Zealand Journal of Marine and Freshwater Research,* 3: 177-179. MPI (Ministry for Primary Industries) (2016) Managing Sediment and E. coli in the Whangarei Harbour Catchment: Summary Report. Prepared for the Ministry for Primary Industries by: Jane White (Ministry for Primary Industries) Adam Daigneault and John Dymond (Landcare Research New Zealand) Malcolm Green,

Chris Palliser, Sandy Elliot and Chris Tanner (National Institute of Water and Atmospheric Research). April 2016.MPI Technical Report No: 2016/56, ISBN No: 978-1-77665-370-6 (online). 33p.

Pratt, D.R., Lohrer, A.M., Pilditch, C.A., Thrush, S.F. (2014) Changes in ecosystem performance across sedimentary gradients in estuaries. *Ecosystems*, 17: 182–194.

Robertson, B.M. and Stevens, L. (2012). New River Estuary. Fine scale monitoring of highly eutrophic arms 2011/12. Report prepared for Environment Southland. 30p.

Robertson, B.M. and Stevens, L.M. (2013). Jacobs River Estuary. Fine scale monitoring of highly eutrophic arms 2012/13. Report prepared for Environment Southland. 29p.

Robertson, B.M., Gillespie, P.A., Asher, R.A., Frisk, S., Keeley, N.B., Hopkins, G.A., Thompson, S.J., Tuckey, B.J. (2002). Estuarine Environmental Assessment and Monitoring: A National Protocol. Part A. Development, Part B. Appendices, and Part C. Application. Prepared for supporting Councils and the Ministry for the Environment, Sustainable Management Fund Contract No. 5096. Part A. 93p. Part B. 159p. Part C. 40p. plus field sheets.

Robertson, B.M., Stevens L.M., Robertson, B.P., Zeldis, J., Green, M., Madarasz-Smith, A., Plew, D., Storey, R., Oliver, M. (2016b). NZ Estuary Trophic Index Screening Tool 2. Determining Monitoring Indicators and Assessing Estuary Trophic State. Prepared for Envirolink Tools Project: Estuarine Trophic Index, MBIE/NIWA Contract No: C01X1420. 68p.

Robertson, B.P., Gardner, J.P.A. & Savage, C. 2015. Macrobenthic–mud relations strengthen the foundation for benthic index development: A case study from shallow, temperate New Zealand estuaries. *Ecological Indicators,* 58, pp.161–174.

Thrush S.F., Hewitt J.E., Cummings V.J., Ellis J.I., Hatton C., Lohrer A.M., Norkko A. (2004) Muddy waters: elevating sediment input to coastal and estuarine habitats. *Frontiers in Ecology and the Environment,* 2: 299-306.

Townsend, M., Lohrer, A.M. (in prep.) ANZECC Guidance for Estuary Sedimentation. *NIWA CLIENT REPORT No: HAM2015-096.* Prepared for Ministry for the Environment.

Townsend, M., Thrush, S.F., Hewitt, J.E., Lohrer, A.M., McCartain, L., (2014) Behavioural changes in the Tellinid bivalve *Macomona liliana* (Iredale, 1915) following exposure to a thin terrigenous sediment deposition event: evidence from time-lapse photography. *Cahier de Biologie,* 55: 475-483.

Woods, P.J., Armitage, P.D. (1997) Biological Effects of Fine Sediment in the Lotic Environment. *Environmental Management*, 21(2): 203-17.

Sediment metals

Although metals/metalloids (herein metals) occur naturally to some extent, their prevalence within estuarine sediments can increase due to human-induced changes in land-use (e.g., agriculture and urban development). Land-derived metals can be flushed into streams/rivers and deposited into estuarine sediments, which act as a sink for contaminants (Robertson et al. 2002). In New Zealand estuaries, the concentrations of different metals in sediments are typically correlated (Robertson et al. 2002).

Sediment metals can provide information regarding the condition or state of estuaries and link to values such as those associated with ecosystem health. At certain concentrations, sediment metals can be toxic to benthic organisms (ANZECC 2000), and benthic organisms can also contribute to the bioaccumulation of metals in estuarine food webs (Robertson et al. 2002). Furthermore, metals can bind with fine sediments, which may cause additional stress to benthic organisms living in muddy areas. Some metals in sediments are also generally well correlated with organo-chlorine contaminants (Hewitt et al. 2014).

There are some issues to consider if using sediment metal concentrations as an attribute for managing upstream effects. Anthropogenic activities not associated with land-runoff transported from rivers/streams are associated with increased metal concentrations in estuaries. For example stormwater is often a source of metals/metalloids, and boats have been highlighted as a source of copper contamination (MPI MEMP). Influence of these impacts/processes on metal concentrations may confound upstream effects. Temporal and spatial variability in sediment metal concentrations has been analysed and trends over time have been able to be identified, however their natural temporal variability has so far not precluded detection of trends (Hewitt et al. 2014). Sediment metals can also be relatively expensive to analyse.

Sampling design

National standards / guidelines and consistency: No national standards. Although sampling designs for collecting sediment metals information vary across the country (i.e., site extent, number of replicates, frequency and sampling time, site representativeness). Many of the intertidal SOE ecological monitoring programmes, follow guidelines described in Robertson et al. (2002).

Although most council-led estuary monitoring programmes focus on intertidal habitats, subtidal sediment samples are also collected in New Zealand estuaries for a variety of reasons (e.g., SOE, consent monitoring). Sampling design for subtidal sample collection varies between programmes.

Potential bottlenecks: Inappropriate sampling design, in terms of spatial extent and number of replicates, selection of non-representative sampling sites, not reflecting the entire gradient of degradation/response to a stressor. Differences in the number of replicates (Bolton-Richie & Lawton in draft) and sampling frequency between some programmes.

Opportunities: To optimize sampling strategy, metals sampling can be aligned with collecting data on other benthic attributes and/or state variables (e.g., sediment quality characteristics, macrofauna, shellfish contaminants).

Caveats and recommendations: Analyse the effects of varying frequency and replication. Understand the rationale behind the location of current sampling sites (Bolton-Richie & Lawton in draft). Many of the current council-led metal contaminant sampling programmes target intertidal habitat (Bolton-Richie & Lawton in draft). Although it is usually a dominant and relatively vulnerable habitat, it may not represent the whole estuary condition in response to upstream effects. Therefore, sampling subtidal sites should also be considered for certain estuaries.

Sampling procedures

National standards / guidelines and consistency: No national standards. Metal contaminant sampling reasonably consistent across the country with many ecological programmes following procedures in Robertson et al. (2002). Subtidal sampling is often conducted using divers to collect cores or from a vessel using a grab.

Potential bottlenecks: Possible differences in the depth of sediment collected for analysis (Bolton-Richie & Lawton in draft), and also in compositing of samples.

Opportunities: Additional samples could be collected and analysed using a different laboratory method in order to obtain data on the comparability of results (see Laboratory analyses section below).

Caveats and recommendations: Standardized sample depth and, possibly, sample compositing. It has been noted that between three and five replicate samples are required to adequately assess concentrations of lead, copper and zinc (cited in Hewitt et al. 2014).

Laboratory analyses

National standards / guidelines and consistency: No national standards. The metals cadmium, chromium, copper, lead, mercury, nickel, and zinc, and the metalloid arsenic are measured by many councils (Bolton-Richie & Lawton in draft). There are differences between which sediment fractions are analysed, the different size fractions include: total, <500 um, and <63um) (Bolton-Richie & Lawton in draft), as well as <2000 um (Hills Laboratory pers. comm.). There can be differences between analytical detection limits (ADL), particularly if either trace or screen methods are used for analysis. There are also other differences in analytical methods, e.g., digestion methods.

Potential bottlenecks: Not all main metals are measured in all programmes and metals are analysed from different sediment grain size fractions (Bolton-Richie & Lawton in draft). There are differences in ADLs and slight differences in analytical methods.

Opportunities: Could collect multiple metal samples and analyse them in different ways (e.g., from different sediment size fractions) to determine comparability between methods.

Caveats and recommendations: Investigate the influence of the analysis of different grain size fractions on the concentration of contaminants (Bolton-Richie & Lawton in draft). Standardise the metals analysed for, as well as the laboratory analysis methods and ADLs.

Computational approaches and metrics derived

National standards / guidelines and consistency: No national standards. ANZECC (2000) report as mg/kg. The EMP mentions that sediment metals contaminant values can be normalized to 100% mud content (Robertson et al. 2002). The EMP also mentions that copper may be used as a surrogate for other metals, although it is recognized that high concentrations may result in the need for follow up analyses of other metals (Robertson et al. 2002).

Caveats and recommendations: Determine the metric to report given the different number of samples collected per site (Bolton-Richie & Lawton in draft). Standardise whether metals data are normalized to mud or not.
Thresholds
There are ANZECC (2000) trigger values based on toxic effects to organisms for metals in sediments. However, there is some evidence to suggest that ecological effects occur at metal values lower than national (e.g., ANZECC 2000) low guidelines (Hewitt et al. 2009; Rodil et al. 2013; Tremblay et al. 2017), which are higher than international effects range-low guidelines (e.g., Long and Morgan, 1990), based on equivalent principles. Auckland Regional Council have also developed Environmental Response Criteria against which sediment metals concentrations can be compared (Auckland Regional Council 2004).
References

ANZECC, 2000. Sediment quality guidelines. In: Australian and New Zealand Guidelines for Fresh and Marine Water Quality 2000 (Volume 2: Aquatic ecosystems - rationale and background information). Australian and New Zealand Environment and Conservation Council and Agriculture and Resource Management Council of Australia and New Zealand, Artarmon.

Auckland Regional Council. 2004. Blueprint for monitoring urban receiving environments. Auckland Regional Council, Technical Publication No. 168 (revised), Auckland.

Bolton-Ritchie L, Lawton R (in draft). Summary of Regional Council metadata collected to inform the coastal ecosystem module of LAWA.

Hewitt, J.E., Anderson, M.J., Hickey, C.W., Kelly, S., Thrush, S.F., 2009. Enhancing the ecological significance of sediment contamination guidelines through integration with community analysis. Environmental science & technology 43, 2118-2123.

Hewitt, J.E., Bell, R., Costello, M., Cummings, V., Currie, K., Ellis, J., Francis, M., Froude, V., Gorman, R., Hall, J., Inglis, G., MacDiarmid, A., Mills, G., Pinkerton, M., Schiel, D., Swales, A., Law, C., McBride, G., Nodder, S., Rowden, A., Smith, M., Thompson, D., Torres, L., Tuck, I., Wing, S. (2014) Development of a National Marine Environment Monitoring Programme (MEMP) for New Zealand.

Long, E.R., Morgan, L.G., 1990. The potential for biological effects of sediments-sorbed contaminants tested in the National Status and Trends Program, NOAA technical memorandum NOS OMA. 52.

Robertson BM, Gillespie PA, Asher RA, Frisk S, Keeley NB, Hopkins GA, Thompson SJ, Tuckey BJ 2002. Estuarine Environmental Assessment and Monitoring: A National Protocol. Part A. Development, Part B. Appendices, and Part C. Application. Prepared for supporting Councils and the Ministry for the Environment, Sustainable Management Fund Contract No. 5096. Part A. 93p. Part B. 159p. Part C. 40p. plus field sheets.

Rodil IF, Lohrer AM, Hewitt JE, Townsend M, Thrush SF, Carbines M 2013. Tracking environmental stress gradients using three biotic integrity indices: Advantages of a locally-developed traits-based approach. Ecological Indicators 34, 560-570.

Tremblay, L.A., Clark, D., Sinner, J., Ellis, J.I., 2017. Integration of community structure data reveals observable effects below sediment guideline thresholds in a large estuary. Environmental Science: Processes & Impacts.

Shellfish metals and other chemical contaminants

Aquatic organisms such as shellfish and fish can accumulate substantial levels of chemical and microbial contaminants when exposed to polluted water and sediment. In the case of microbial contamination, this can lead to these organisms being unfit for human consumption (see Factsheet 'Shellfish faecal indicator bacteria (FIB)'). While chemical contamination is not generally high enough in estuarine and coastal areas to be a significant general concern for human consumers of fish or shellfish (Stewart et al. 2013) (exceptions are natural marine biotoxins and some localised contaminated areas), chronic health effects on the aquatic organisms themselves, or on other animals that feed on them, are possible ecological consequences. The primary Contaminants of Potential Concern (COPC) for inorganics are cadmium – from agricultural fertilisers; mercury (Hg) – primarily from natural geothermal sources, copper – from stormwaters and antifouling use, and zinc – from stormwaters. Only mercury biomagnifies in the food-chain, with tissue concentrations increasing with each higher trophic level because of methyl-mercury formation and accumulation in fatty tissue. Some persistent bioaccumulative and toxic (PBT) organic chemicals, such as organochlorine pesticides (legacy agricultural inputs of dieldrin and DDT) and industrial polychlorinated biphenyls (PCBs), can significantly bioaccumulate in the tissues of some aquatic organisms and may cause chronic, long-term ecological problems. A wide range of other organic chemicals are discharged to the coastal marine environment (e.g., agriculture/forestry herbicides and pesticides; industrial chemicals; personal care products; and endocrine disrupting chemicals, such as human contraceptives). Many of these organic chemicals do not markedly bioaccumulate in the tissues of aquatic species (being hydrophilic or metabolized), but can be detected by chemical analyses of biota tissues. Biological monitoring is often useful in providing a cumulative assessment of chemical contaminants – many of which occur in highly time-varying exposures, such as for stormwater discharges. Species that are likely to accumulate highest levels of contaminants are those that live in contaminated environments, particularly when exposed to or feeding on polluted sediments e.g., shellfish, snails, bottom-feeding fish, and worms.

Some chemicals are transferred through the food chain, so higher trophic level organisms, in particular birds, that feed on contaminated worms, fish, and shellfish, can accumulate high concentrations, and this can cause serious ecological problems (e.g., the infamous egg-shell thinning problems for American birds of prey, caused by exposure to organochlorine pesticides such as DDT). Human health risks from bioaccumulation are highly significant in other countries, with the USA National Coastal Condition Report (US EPA 2016), finding that 77% of sites throughout the coastal USA have 'unsatisfactory' fish tissue concentrations, mainly due to PCBs, Hg and DDT.

The major standardized marine shellfish biomonitoring programme for contaminants is the 'Mussel Watch' programme run by NOAA in the United States, with over three decades of monitoring and an analyte list that has grown to over 140 chemicals (Kimbrough et al. 2008). This mussel watch programme has been used to detect long-term trends and emerging contaminants present in US coastal waters (Melwani et al. 2014). Internationally, various mussel watch monitoring programmes have been established in European and Asian marine ecosystems. In New Zealand, monitoring of chemical contaminants in shellfish is closely linked to food safety, and therefore often falls under the jurisdiction of public health (Ministry of Health (MoH)) or commercial shellfish export certification (Ministry of Primary Industries (MPI)) programmes. The Auckland Regional Council (ARC, now Auckland Council) operated a Shellfish Contaminant Monitoring Programme for metals and organic contaminants originating with oysters in 1987 and mussels in 1999 (ARC 2007), until its closure in 2016 because of costs (M. Cameron, Auckland Council, pers com). The Bay of Plenty Regional Council

have undertaken an oyster monitoring programme for contaminants in Tauranga Harbour since 1990. Presently there is no standard nationwide shellfish monitoring programme operating in New Zealand. Developing shellfish biomonitoring-based attributes might be complicated by the variability of contaminant exposure conditions in estuaries (i.e., internal vs external sources) and the suitability of local reference (benchmarking) sites. This may impede establishing stressor-specific "health" responses and setting contaminant-specific bands which are linked to upstream management.
Sampling design
National standards / guidelines and consistency: Shellfish biomonitoring can be undertaken using wild *in situ* 'sentinel shellfish' or caged shellfish. The choice of the approach used will be dependent on the objectives of the monitoring programme. The use of caged shellfish provides a more standardised approach to the biomonitoring programme. The source, size, numbers and biological condition of the shellfish deployed at each site can be standardised. Additionally, caged shellfish can be located at appropriate monitoring sites, including reference sites and surveillance monitoring sites. The surveillance monitoring sites can include stations designed for long-term trend monitoring, discharge-specific effects (e.g., municipal, industrial or stormwater discharges) or land-use dominated effects (e.g., urban, rural). Caged shellfish can be deployed at standardised depths in the water column – thus resulting in comparable exposures to tidal and wind-induced bed sediments. Optimising a monitoring programme for land-use related effects would ideally use a hydrodynamic model of the estuary to assist in characterising the location of monitoring sites in relation to contaminant sources within and external to the estuary. This would aid in establishing anticipated contaminant exposure gradients and reduce the potential for redundancy in site selection. Physiological tolerances for salinity, water temperature and species distributions should be considered in the design of the programme for choice of shellfish species. For example, the Auckland Council shellfish biomonitoring programme used both mussels and oysters – with oysters being deployed at sites experiencing a wide range in salinity from freshwater inflows from river inputs (Stewart et al. 2013). Management considerations for the design of a shellfish biomonitoring programme include: **Values:** Human health, Ecosystem health, Mahinga kai. **Species selection:** Mussels (for high salinity environments); Oysters (variable salinity), Mahinga kai applications – various (species 'food basket' may be required for suit of species consumed; e.g., cockles, horse mussels, paua, pipi, scallops, tuatua). **Duration:** 2-3 months generally required to allow chemical exposure and uptake. Physiological response will also require significant exposure period to establish body-burden and effects. **Frequency:** Programme may have a monitoring frequency which ranges from annual to 5-yearly. The design of the monitoring programme should consider changes occurring in the estuary catchment and resourcing requirements in determining the monitoring frequency. A tiered approach to the study design could include long-term "trend" sites and "impact" sites in addition to the reference sites.

Frequency of monitoring and objectives, including the suite of analytes, could differ between the various types of monitoring sites.

Adverse effects assessment: Condition (dry weight/shell volume); Energy stores (glycogen, lipids); Biomarker responses to chemicals (e.g., metallothionein); Genetic measures (various).

Tissue archiving: Sub-samples of shellfish tissue can be archived for future analyses of other chemicals of concern. Reference tissues can also be archived to provide for reference benchmarking should future analytical methods change.

Other measures: Shell can also be analysed to for chemical contaminants. Ablation methods can be used to provide a time series for contaminant exposures for resident shellfish.

Quality control: International reference tissues are available for use as part of the analytical quality control procedures for specific chemical classes.

Special applications: Food basket surveys for the "total dietary survey" (NZFSA 2004) and for non-commercial wild food (NZFSA 2005) and site-specific studies (Whyte et al. 2009; Phillips et al. 2014). Some outfall monitoring programmes incorporate use of mussels deployed along transects for discharge consent monitoring – including chemical contaminants and microbial indicators. Physiological response studies of shellfish to marine toxins have been undertaken for surf-clams and green-lipped mussels (Marsden et al. 2015).

Disadvantages:

1. Caged mussels may be subjected to vandalism. Appropriate methods needed to ensure low visibility of cages.
2. Treatments may be lost in extreme weather and through marine activities (e.g., trawling).
3. Measures of adverse effects (e.g., condition reduction) will always represent cumulative effects of all contaminant exposures – and potential differences in food supply. Reference sites and deployment along exposure gradients can assist in differentiating effects attributable to specific sources.
4. Limited number of contaminants with human health standards. Attribute bands would need to be based on tissue chemical analysis and difference from reference (i.e., benchmark) sites.

Potential bottlenecks: Shellfish biomonitoring data is limited in terms of spatial extent and number of replicates, selection of surveillance and reference sampling sites. If a long-term monitoring programme is implemented (e.g., 5-yearly), then a long period will be required before statistical data (e.g., median, 95th percentile) can be collected for between site comparisons, or for statistical detection of trends.

Opportunities: Shellfish are a good integrator of contaminant variability in the water column, and with appropriate sampling design could be used as a means of tracking trends in both water and shellfish quality. Shellfish biomonitoring programme for chemical contaminants can be combined with faecal indicator bacteria (FIB) in shellfish. Combined programme will provide greater information for human health risk assessment.

<u>**Caveats and recommendations**</u>: Site selection for surveillance and reference monitoring sites, and species selection, are critical to the success of the monitoring programme. Specific-site locations and knowledge of estuary hydrodynamics are required to link shellfish contaminants with local catchment inputs. Some chemical contaminants of potential concern may have limited bioaccumulation. Analytical detection ability may be limited by tissue extraction and analytical clean-up requirements, together with analytical method detection limits. Specific high sensitivity methods may be applicable to specific contaminants of concern. Shellfish tissue can be archived for future chemical analysis of contaminants.
Sampling procedures
<u>**National standards / guidelines and consistency**</u>: No national standards. Protocols for caged mussel and oyster monitoring are available from existing long-term programmes (Stewart et al. 2013; Park 2016). Protocols for field deployment of other shellfish species (e.g., cockles, wedge shell) would need to be developed. Spatial design of monitoring programme and criteria for reference sites requires standardization. <u>**Potential bottlenecks:**</u> Bottlenecks include those as outlined under sampling design and relate to the timing (i.e., season) and conditions when deployments are undertaken, and insufficient spatial and temporal replication of samples. <u>**Opportunities:**</u> Monitoring programmes can include both resident and caged species – potentially including a range of species for a 'food-basket' approach for Mahinga kai monitoring. Species with different feeding modes can help distinguish water and sediment pathways for contaminants entering food-chain. <u>**Caveats and recommendations**</u>: A standardised size range and deployment period will improve quality control and inter-comparison between estuaries in different regions.
Laboratory analyses
<u>**National standards / guidelines and consistency:**</u> No national standards. Chemical analyses for tissue and shell are available for a wide range of common chemical contaminants from commercial laboratories (e.g., Hill Laboratories Ltd, Hamilton; AssureQuality, organic chemicals). Specialist analytical laboratories are generally required for trace organic compounds. The US mussel watch programme currently has a suite of 140 compounds for analysis using standard and developing methods (Kimbrough et al. 2008). Standard tissues are available to include in the analytical QC procedures. General measures of shellfish 'health' (e.g., condition, glycogen content) are not available from commercial laboratories. Total lipid content is generally available from laboratories undertaking organic analyses. Specific biomarker measures (e.g., metallothionein) are not available from commercial laboratories.

<u>Potential bottlenecks:</u> Chemical contaminant detection limits may differ between analytical laboratories. Contaminants of concern may not be analysed in standard suite of chemical analyses – requiring specialist laboratory services or development of new techniques. Tissue can be archived (generally frozen at -80°C) for subsequent analyses. **<u>Opportunities:</u>** Chemical contaminant biomonitoring can be combined with monitoring programme for shellfish FIB. However, the time scales of the two monitoring programmes will differ – requiring longer term exposure for chemical contaminant bioaccumulation. Shellfish samples collected during and immediately following periods when FIB contamination is greatest (e.g., periods of flooding) are required to develop relationships between levels of contamination and upstream contaminant loading. **<u>Caveats and recommendations</u>**: Robust QA procedures should be incorporated into the chemical analytical procedures. Archived samples should be retained from key sites for potential analysis for additional 'new' contaminants once suitable analytical methods are available, or for reanalysis of existing contaminants using updated methods with improved detection limits.
Computational approaches and metrics derived
<u>National standards / guidelines and consistency:</u> Limited to the narrative standards in RMA and the standardised thresholds for human consumption set out below. Quantitative values for limits to dietary consumption quantity and frequency may be calculated for a wide range of chemical contaminants (Stewart et al. 2011; Phillips et al. 2014). Comparative data from reference sites can be used for common contaminants (e.g., copper, zinc) and physiological data (e.g., condition) to establish effects bands – which would be species-specific. **<u>Potential bottlenecks:</u>** Bioaccumulation and health effects metrics derived based on comparative data with reference sites will be species-specific. Comparisons between sites and over time may be limited by analytical detection limits – which may differ between laboratories and over time as methods change. Analytical detection limits for key contaminants of concern need to be sufficiently low to detect chemical exposures at concentrations which are environmentally relevant to potential adverse effects on aquatic organism. Need statistical and reporting methods that deal with data measurements less than the detection limits. Multiple estuarine monitoring sites and a spatial component of the monitoring design will be required for establishment of local reference sites and distinguishing catchment from local legacy and harbour-generated sources. **<u>Opportunities:</u>** Multiple metrics of exposure and 'health' effects on organisms can be used. **<u>Caveats and recommendations</u>**: Any measures of shellfish 'health' will represent the composite exposure effects of all chemical contaminants – combined with potential food quantity/quality effects on growth. Seasonal reproductive cycles will also significantly affect contaminant body-burdens – particularly of organic contaminants – and health measures, such as condition and glycogen/lipid content. Standard monitoring programmes for trend and effects detection should be undertaken over comparable seasonal periods which avoid reproductive periods. Durations of exposure of caged shellfish need to be standardized and sufficiently long to allow bioaccumulation and for physiological response to contaminant body-burdens (e.g., 3 months duration). Compositing of shellfish from a site may be undertaken prior to undertaking chemical analyses to reduce biological variability.

Thresholds (existing criteria)

The RMA Schedule 3 specifies a narrative bioaccumulation standard for shellfish gathering (SG) waters, and they are probably implicit in Class AE and C Waters (Table 1).

Table 1 - Standards for Bioaccumulation for Water Quality Classes from Schedule 3 RMA

Class	Purpose	Criteria
AE	Aquatic Ecosystem	The following shall not be allowed if they have an adverse effect on aquatic life: (c) any discharge of a contaminant into water.
SG	Gathering/Cultivating Shellfish	Aquatic organisms shall not be rendered unsuitable for human consumption by the presence of contaminants.
CR	Contact Recreation	-
IA	Industrial Abstraction	-
NS	Natural State	The natural quality of the water shall not be altered.
A	Aesthetic	
C	Cultural	The quality of the water shall not be altered in those characteristics which have a direct bearing upon the specified cultural or spiritual values.

Food safety

The Australia New Zealand Food Standards Code (Food Standards Australia and New Zealand 2015) prescribes maximum levels for arsenic (As), cadmium (Cd), lead (Pb), mercury (Hg), polychlorinated biphenyls (PCB), histamine and marine biotoxins in seafoods. Standard 1.4.1, Contaminants and natural toxicants, sets out the maximum levels (MLs) of specified metal and non-metal contaminants and natural toxicants in nominated foods (http://www.foodstandards.gov.au/foodstandards/foodstandardscode.cfm).

The maximum levels for the toxic metals are summarised in Table 2.

Table 2 - Food standards (mg/kg, wet weight) for Australia and New Zealand (Food Standards Australia and New Zealand 2015).

Contaminant	Crustaceans	Fish	Molluscs	Seaweed (edible kelp)
Arsenic (As) [a]	2	2	1	1
Cadmium (Cd)			2 [b]	
Lead (Pb)		0.5	2	
Mercury (Hg) [c]	0.5	0.5-1.0	0.5	

[a] Value based on inorganic arsenic; [b] Excluding dredge/bluff oysters and queen scallops; [c] Two separate maximum levels are imposed for fish — a level of 1.0 mg mercury/kg (as a mean) for the fish that are known to contain high levels of mercury (such as long-lived or large marine species) and a mean level of 0.5 mg/kg for all other species of fish. A mean limit of 0.5 mg/kg is also imposed for crustacea and molluscs. The Australia New Zealand Food Standards Code also specifies a standard

based on the number of serves (meals) of different fish that can be safely consumed (Food Standards Australia and New Zealand 2015). For the other common heavy metals Cr, Cu, Ni and Zn, Turner et al. (2005) state "*These four heavy metals are environmentally ubiquitous in New Zealand, and their levels are often higher in areas associated with human activity. For this reason, they are commonly included for analysis during heavy metal studies. While toxic to humans at high concentrations, Cu, Zn and probably Cr are essential elements and all are well regulated by the body. For this reason, their concentrations in foods are not regulated by the NZFSA and there are no food safety limits in New Zealand*." *Ecological effects* There are no criteria to protect aquatic animals from bioaccumulation and biomagnification effects.
Critique/review of existing approaches
Human health Internationally, concentration levels that protect human consumers have decreased in food for some organochlorines in recent years and there has been an increased awareness that some members of the community consume or wish to consume larger and more frequent meals of seafood, which may include parts of fish which bioaccumulate more contaminants (the standard fish advisory assumes consumption of fillets). The Australia New Zealand Food Standards Code prescribes maximum levels for As, Cd, Pb, Hg, PCB, histamine and marine biotoxins in seafoods. The consumption standards are based on a life-time consumption of a standard dietary intake. Some oysters are allowed to breach the Cd standard by large amounts, because levels are regarded as "natural". The standards do not address DDT or dioxins and furans, which have been found to trigger advisory notices around fish consumption in the USA. Maximum consumption recommendations need to be specifically developed for members of the population that aspire to consume a wider variety of fish and shellfish and/or regularly have a higher dietary intake collected from coastal water bodies. To do this, a full Health Risk Assessment (HRA) for food consumption relevant to Māori and other ethnic groups would need to be undertaken. This would involve measuring Hg, Pb, Cd, As, PCB, dioxins and DDT levels in targeted species and assessing the health risk associated with a "food basket" of the same widely utilised species, and include an additive risk assessment for multiple contaminants (Stewart et al. 2011; Phillips et al. 2014). Concentrations of PCBs and DDTs, while not especially high in terms of toxicity effects, could probably trigger bioavailability studies or even fish advisories in the USA, and would be expected to markedly reduce the recommended levels of dietary intake for many species (Stewart et al. 2011). HRA could also consider differing risk categories (general population, women of child-bearing age, children) and for realistic levels of consumption (moderate and high consumers), or utilise guidance from the most sensitive for establishing the "guidelines". Outcomes may be no risk from "normal" consumption levels or the need for guidance to limit consumption. The application and methodology have been developed in the Bay of Plenty region (Phillips et al. 2014) and are proposed for the Waikato River clean up (NIWA 2010). This type of monitoring could be applied to areas identified and classified/zoned for gathering/cultivating shellfish. *Ecological health*

Biomarkers show biochemical and/or physiological changes in an organism following exposure to contaminants. Various biomarkers could be used as basis for determining thresholds of adverse effects because of chemical contaminant exposure. For shellfish, measures of condition (i.e., tissue mass/shell volume) and total energy reserves (e.g., glycogen or lipid content) are general non-specific 'health' measures that can differentiate sites based on contaminant body-burdens (Roper et al. 1991; Hickey et al. 1995). Non-specific biomarkers may indicate that the biomonitoring organisms have been exposed to a toxicant/stressor, but the response is not necessarily related directly to a toxicity-specific mechanism, with factors such as reduced food abundance/quality also potentially contribution to a low condition state. Various other biochemical and specific physiological measures (e.g., metallothionein protein response to heavy metals) have been used on shellfish in New Zealand estuaries (see earlier section) – providing techniques which are suitable for determining causation linkages with specific contaminant classes. They can be very sensitive indicators of sub-lethal ecological effects and provide both quantitative and qualitative estimates of exposure (van der Oost et al. 2003). Biomarker thresholds for physiological change will differ with the type of chemical contaminant and between species. Additionally, there will be seasonal changes in organism 'health' in relation to food supply, natural reproductive cycles and potentially through extreme events resulting in habitat disturbance. Because of the range of potential contaminant effects and species-specific differences in sensitivity, effect thresholds relating to adverse effects are recommended to be based on comparisons of sites with a local reference site. The conditions at the local reference site should be representative of high water quality by being distant from known contaminant sources, but incorporate changes in food supply and coastal salinity which as close as practicably represent the conditions at the key monitoring sites. Species selection. The species most commonly used in New Zealand for shellfish biomonitoring are oysters (*Crassostrea gigas*) and mussels (green shell, *Perna canaliculus*). As the distribution of the green shell mussel is limited to the upper South Island northwards, the use of the blue mussel (*Mytilus galloprovincialis*) would be considered the standard mussel species for southern New Zealand waters. The cockle (*Austrovenus stutchburyi*) is a water column filtering species and the wedge shell (*Macomona liliana*) is a deposit-feeding species. Field and laboratory biomonitoring studies have been undertaken for organic contaminants (Hickey et al. 1995) and metals (Purchase and Fergusson 1986; Fukunaga and Anderson 2011). Generally, the deposit-feeding wedge shell shows markedly higher bioaccumulation of organic and metal contaminants than does the filter-feeding cockle. Deployment period. Durations of exposure of caged shellfish need to be standardized and sufficiently long to allow bioaccumulation and for physiological response to contaminant body-burdens (e.g., 3 months duration).
Relevance/suitability for national application
Auckland Council's monitoring of resident and deployed shellfish show that chemical contaminants, Zn, Cu and Pb and organic compounds including PAH, OCPs, and PCBs, are accumulated by these biota from the water column, enabling spatial patterns and temporal trends in contamination to be measured. This programme has previously been an important part of Auckland Council's state of the environment monitoring. By international standards, organic contaminant concentrations in mussel and oyster tissues are low and are unlikely to cause ecological or health effects.

There have been a number of other research studies in Auckland that have measured bioaccumulation, and these have been reviewed by Kelly (2009). However there has been little or no assessment of effects on animals, ecology or human consumers. Of these, the only identifiable effects of bioaccumulation are Pb levels in oyster catchers from Mangere Inlet which might induce chronic toxicity (Thompson and Dowding 1999). What can be concluded from all these studies is that bioaccumulation can occur with priority contaminants and, as expected, this is consistent with overseas studies, although concentrations in Auckland are generally much lower. There are some indications of potential toxicity to higher animals (oyster catchers, flounder), and while the evidence is not strong, because they are preliminary studies only, ecological effects from bioaccumulation cannot be ruled out.
Conclusions
We are unable to recommend criteria for bioaccumulation/ biomagnification to protect aquatic life. However, we recommend a review/study of the situation for legacy contaminants Hg, PCB and DDT accumulation in the local food chain, in order to assess the risk of these contaminants to higher animals, especially human consumers and New Zealand threatened and endangered birds. In general, measurement of contaminants in biota may yield some useful information as to whether or not a contaminant is bioavailable. However, such studies need to be conducted skilfully because some contaminants may be bioavailable and toxic, but not bioaccumulate, while some animals may regulate and minimise the bioaccumulation of a toxic chemical. Bioaccumulation is an important component of special investigations into the fate and effects of bioaccumulative toxic contaminants, such as in toxicity studies or in Weight of Evidence approaches. In terms of human consumers, there are few reports of high risks to human health from accumulation of priority pollutants in aquatic organisms in New Zealand, except for mercury, as noted above. Cadmium levels exceed food safety limits in oysters, but this seems to be a natural phenomenon. This situation could be worsened by the build-up, and subsequent runoff of Cd in pasture soils from superphosphate application (Butler and Timperley 1996). Both Cd and uranium are elevated in soils as a result of phosphatic fertiliser additions (Taylor 2007; Schipper et al. 2011; Salmanzadeh et al. 2017) – though we are not aware of any marine biomonitoring studies which have measured tissue uranium concentrations. The recommended guideline values for bioaccumulation protection of human health values for priority pollutant toxic contaminants are summarised in Table 3. Other guidelines for aquatic ecosystem and natural state protection will need to be derived on a reference site approach. This would include major stormwater and land-use derived metal contaminants – such as copper, zinc and uranium – together with ‘health’ measures for biomonitoring species.

Table 3 - Recommended Guidelines for Bioaccumulation protection of Human Health Values for Priority Pollutant Toxic Contaminants

Class	Purpose	Criteria
AE	Aquatic Ecosystem	-
SG	Gathering/Cultivating Shellfish	Food Standards Australia and New Zealand (2015) for As, Cd, Pb, Hg, PCB Food Standards Australia and New Zealand (2015) limits of consumption of types of fish and sensitive members of the population
CR	Contact Recreation	-
IA	Industrial Abstraction	-
NS	Natural State	-
A	Aesthetic	-
C	Cultural	Develop food basket approach to assess and manage risk from Hg, PCB, DDT, As, Cd, Pb and dioxins in seafoods

References

ARC (2007). Contaminant monitoring in shellfish: results of the 2005 shellfish contaminant monitoring programme. No. ARC Technical Publication TP 332, produced by S. Kelly. Auckland Regional Council, Auckland. pp. 76.

Butler, C.A.; Timperley, M.H. (1996). Fertilised farmland as a source of cadmium in oysters. *The Science of the Total Environment 181*: 31-44.

Food Standards Australia and New Zealand (2015). Australia, New Zealand food standards code - Standard 1.4.1 - Contaminants and Natural Toxicants (http://www.foodstandards.gov.au/code/Pages/default.aspx). www.foodstandards.gov.au.

Fukunaga, A.; Anderson, M.J. (2011). Bioaccumulation of copper, lead and zinc by the bivalves *Macomona liliana* and *Austrovenus stutchburyi*. *Journal of Experimental Marine Biology and Ecology 396*: 244-252.

Hickey, C.W.; Roper, D.S.; Holland, P.T.; Trower, T.M. (1995). Accumulation of organic contaminants in two sediment dwelling shellfish with contrasting feeding modes: deposit - (*Macomona liliana*) and filter-feeding (*Austrovenus stutchburyi*). *Archives of Environmental Contamination and Toxicology 29*: 221-231.

Kelly, S. (2009). Effects of stormwater on aquatic ecology in the Auckland region: A literature review. No. Technical Report TR2010/021. Auckland Regional Council, pp. 65.

Kimbrough, K.L.; Lauenstein, G.; Christensen, J.; Apeti, D. (2008). An assessment of two decades of contaminant monitoring in the Nation's Coastal Zone. No. NOAA Technical Memorandum NOS NCCOS, 74. NOAA/National Centers for Coastal Ocean Science, Silver Spring, MD. pp. 118.

Marsden, I.D.; Contreras, A.M.; MacKenzie, L.; Munro, M.H.G. (2015). A comparison of the physiological responses, behaviour and biotransformation of paralytic shellfish poisoning toxins in a surf-clam (*Paphies donacina*) and the green-lipped mussel (*Perna canaliculus*). *Marine and Freshwater Research 67*: 1163.

Melwani, A.R.; Gregorio, D.; Jin, Y.; Stephenson, M.; Ichikawa, G.; Siegel, E.; Crane, D.; Lauenstein, G.; Davis, J.A. (2014). Mussel watch update: Long-term trends in selected contaminants from coastal California, 1977-2010. *Marine Pollution Bulletin 81*: 291-302.

NIWA (2010). Waikato River Independent Scoping Study (WRISS). No. MFE10201; HAM2010-032. NIWA report by J.C. Rutherford and J.M. Quinn (and others) for the Ministry for the Environment, Hamilton. pp. 273 + App.

NZFSA (2004). New Zealand Total Diet Survey. http://www.nzfsa.govt.nz/science/research-projects/total-diet-survey/. Wellington, Institute of Environmental Science & Research report by R.W. Vannoort and B.M. Thomson for New Zealand Food Safety Authority.

NZFSA (2005). Review of Non-Commercial Wild Food in New Zealand. No. NZFSA Draft Position Paper (No: 01/05). New Zealand Food Safety Authority, Wellington. pp. 28.

Park, S. (2016). Comparison of contaminant levels in oysters from Tauranga Harbour in 1990 and 2016. No. Environmental Publication 2016/09. Bay of Plenty Regional Council, pp. 43.

Phillips, N.; Stewart, M.; Olsen, G.; Hickey, C.W. (2014). Human health risks of geothermally-derived metals and other contaminants in wild-caught food. *Journal of Toxicology and Environmental Health, Part A: Current Issues 77*: 1-20.

Purchase, N.G.; Fergusson, J.E. (1986). Chione (*Austrovenus stutchburyi*), a New Zealand cockle, as a bio-indicator for lead pollution. *Environmental Pollution 11*: 137-151.

Roper, D.S.; Pridmore, R.D.; Cummings, V.J.; Hewitt, J.E. (1991). Pollution related differences in the condition cycles of Pacific Oysters *Crassostrea gigas* from Manukau Harbour, New Zealand. *Marine Environmental Research 31*: 197-214.

Salmanzadeh, M.; Schipper, L.A.; Balks, M.R.; Hartland, A.; Mudge, P.L.; Littler, R. (2017). The effect of irrigation on cadmium, uranium, and phosphorus contents in agricultural soils. *Agriculture Ecosystems & Environment 247*: 84-90.

Schipper, L.A.; Sparling, G.P.; Fisk, L.M.; Dodd, M.B.; Power, I.L.; Littler, R.A. (2011). Rates of accumulation of cadmium and uranium in a New Zealand hill farm soil as a result of long-term use of phosphate fertilizer. *Agriculture Ecosystems & Environment 144*: 95-101.

Stewart, M.; Gadd, J.; Ballantine, D.J.; Olsen, G. (2013). Shellfish contaminant monitoring programme: Status and trends Analysis 1987 - 2011. No. Auckland Council, pp. 208.

Stewart, M.; Olsen, G.; Gadd, J. (2013). Shellfish contaminant monitoring programme: review. No. Auckland Council, pp. 96.

Stewart, M.; Phillips, N.; Olsen, G.; Hickey, C.W.; Tipa, G. (2011). Organochlorines and Heavy Metals in Wild Caught Food as a Potential Human Health Risk to the Indigenous Māori Population of South Canterbury, New Zealand. *Science of the Total Environment 409*: 2029-2039.

Taylor, M.D. (2007). Accumulation of uranium in soils from impurities in phosphate fertilisers. Landbauforschung Völkenrode 2. *57***:** 133-139.

Thompson, D.R.; Dowding, J.E. (1999). Site-specific heavy metal concentrations in blood of South Island pied oystercatchers *Haematopus ostralegus finschi* from the Auckland Region, New Zealand. *Marine Pollution Bulletin 38*: 202-206.

Turner, N.; Cressey, P.; Lake, R.; Whyte, R. (2005). Review of non-commercial wild food in New Zealand. No. Environmental Science & Research Ltd report for New Zealand Food Safety Authority, Wellington. pp. 181.

US EPA (2016). National Coastal Condition Assessment 2010 (http://www.epa.gov/national-aquatic-resource-surveys/ncca). No. EPA 841-R-15-006. U.S. Environmental, Office of Water, Office of Research and Development, Washington, DC. pp. 129.

van der Oost, R.; Beyer, J.; Vermeulen, N.P.E. (2003). Fish bioaccumulation and biomarkers in environmental risk assessment: a review. *Environmental Toxicology and Pharmacology 13*: 57-149.

Whyte, A.L.H.; Raumati Hook, G.; Greening, G.E.; Gibbs-Smith, E.; Gardner, J.P.A. (2009). Human dietary exposure to heavy metals via the consumption of greenshell mussels (*Perna canaliculus* Gmelin 1791) from the Bay of Islands, northern New Zealand. *Science of the Total Environment 407*: 4348-4355.

Shellfish faecal indicator bacteria (FIB)

Faecal indicator bacteria (FIB), bacteria present in the gut of animals and excreted with faeces, can provide an indication of faecal contamination of shellfish that are farmed commercially or harvested for recreational and customary purposes. Common species monitored for shellfish FIB include Greenshell™ mussels, oysters and cockles. Monitoring of FIB in shellfish is closely linked to food safety, and therefore often falls under the jurisdiction of public health and commercial shellfish sanitation programmes. Commercial shellfish harvest regulations are based on the abundance of the faecal indicator bacteria *Escherichia coli* within shellfish tissues. Closures of shellfish areas are managed in a conservative manner and are typically governed by water quality proxies for *E. coli* contamination, such as river flows, salinity and turbidity. The testing of *E. coli* concentrations in shellfish during and post-harvest, and in many cases pathogens (e.g., norovirus), are then used to confirm levels meet export standards. These tests are typically completed during times that are unlikely to result in exceedances of standards (i.e., when harvest areas are open), thereby limiting the amount of data available for developing relationships between shellfish contamination and upstream contaminant loads. The MfE/MoH (2003) Guidelines includes advice for recreational shellfish gathering and recommends monitoring of faecal coliforms in water for assessing quality of shellfish. Faecal coliforms are less specific to humans than *E. coli* and enterococci but are considered more suitable for general assessments of faecal contamination in shellfish gathering water (MfE/MoH 2003).
Sampling design
National standards / guidelines and consistency: Guidelines for recreational water quality sampling, which is intended to encompass contact recreation for both bathing and shellfish harvesting are set out at http://www.mfe.govt.nz/publications/international-environmental-agreements/microbiological-water-quality-guidelines-marine#notehi. Beyond general recommendations (e.g., around sampling frequency), there appears to be no standard or consistent sampling design for monitoring the quality of recreational and customary shellfish and harvest areas. Some outfall monitoring programmes incorporate use of mussels deployed along transects for discharge consent monitoring. A good example is the on-going Bell's Island Outfall monitoring within and outside Waimea Estuary in the Tasman District. These are designed for specific projects, as opposed to following a standardised sampling design. Before a new growing area for commercial harvest can be classified and listed, a sanitary survey is conducted by an MPI officer. This includes a survey, water and shellfish flesh studies, and development of a management plan outlining how risks to shellfish quality will be mitigated. Growing areas are maintained through ongoing sampling of shellfish and the surrounding water. The scope and frequency of sampling is set out in the management plan for each area; hence each sampling programme is site specific as opposed to being standardised across all sites. **Potential bottlenecks:** Shellfish FIB data is limited in terms of spatial extent and number of replicates, selection of non-representative sampling sites. Samples tend to be biased toward good weather conditions and when shellfish are being harvested, rather than being collected during periods when contamination is occurring (e.g., following high rainfall, mobilising sediments). As a result, it will be difficult to develop relationships between upstream loading of faecal contaminants and FIB in shellfish based on the data available. As outlined in Milne et al. (2017), varying approaches

have been adopted for reporting microbial water quality state, such as the minimum number of sample results, the length of season, and which sample statistic(s) to use (e.g., median, 90th percentile, 95th percentile). Nevertheless, the MfE/MoH (2003) guidelines do contain some advice on sampling requirements, viz. "Sampling to test compliance shall be over the whole shellfish-gathering season. A sufficient number of samples should be gathered throughout the gathering season to provide reasonable statistical power in testing for compliance for both the median limit and the 90% samples limit".

Opportunities: Shellfish are likely a good integrator of FIB variability in the water column, and with appropriate sampling design could be used as a means of tracking trends in both water and shellfish quality. Shellfish FIB will be useful as a state variable, and in sampling campaigns could be used to ground-truth predictive, operational models for developing relationships between estimates of potential FIB concentrations in shellfish and levels of upstream faecal contaminant loading. If accessible, there may be value in combining and analyzing results from multiple sanitation surveys for commercial growing areas that lie within estuaries.

Caveats and recommendations: Several factors need to be considered when collecting water for FIB tests, and when interpreting results. As outlined in Green and Cornelisen (2013), risk of faecal contamination varies according to surrounding catchments and land use and the hydrological characteristics of the coastal water body (e.g., flushing). Wave action, climate and water depth also influence FIB concentrations as the bacteria are known to persist in sediments and beach sands and may spike without recent rainfall. Bacteria and viruses are also more prevalent in turbid waters where microbes attach to particles that prolong survival due to solar shading and extend microbe transport distance. As a result, there are a number of parameters that may influence levels of microbial contamination (elevated FIB), including rainfall, solar radiation, tidal state, water clarity and suspended sediments (or turbidity), light penetration, salinity and water temperature. Due to high variability, modelled estimates for shellfish FIB concentrations in response to upstream loading based on land uses and varying conditions may be required to develop a shellfish FIB attribute, whereas measured concentrations could be used for supporting an attribute and as a state variable.

Sampling procedures

National standards / guidelines and consistency: Sampling protocols to measure recreational water quality for contact recreation relating to bathing and shellfish harvest are set out at http://www.mfe.govt.nz/publications/international-environmental-agreements/microbiological-water-quality-guidelines-marine#notehi. Also, the MfE/MoH guidelines state that "The MPN method as described in Standard Methods for the Examination of Water and Wastewater; American Public Health Association (current edition), must be used to enumerate faecal coliforms unless an alternative method is validated to give equivalent results for the waters being tested."

Potential bottlenecks: Bottlenecks include those as outlined under sampling design and relate to the timing and conditions when samples are collected, and insufficient spatial and temporal replication of samples.

Laboratory analyses
<u>National standards / guidelines and consistency:</u> National standards and consistency: For evaluating faecal coliform bacteria concentrations in water, Membrane filtration (APHA 9222D) and Multiple tube (APHA 9221E) measurement procedures are being used. Both procedures presumed to give comparable results. The MPI advocated method for enumerating *E. coli* in shellfish "is a two-stage, five-tube three-dilution most probable number (MPN) method. The first stage of the method is a resuscitation step requiring inoculation of minerals modified glutamate broth (MMGB) with a series of diluted sample homogenates and incubation at 37±1°C for 24±2 hours. The presence of E. coli is subsequently confirmed by subculturing acid producing tubes onto agar containing 5-bromo-4-chloro-3-indolyl-β-D glucuronide and detecting β -glucuronidase activity after incubation (MPI 2013; Enumeration of Escherichia coli in Bivalve Molluscan Shellfish MPI Method Version 9). **<u>Potential bottlenecks:</u>** Culture based methods require at least 24 hours incubation, hence cannot be used to manage harvest areas in "real time". Data is limited for samples collected during and immediately following periods when contamination is greatest (e.g., periods of flooding); this in turn can limit ability to develop relationships between levels of contamination and upstream contaminant loading. **<u>Opportunities:</u>** A faster method of assessing FIB would reduce alarm fatigue and improve compliance with the warnings. Potential use of molecular markers for targeting pathogens and source-specific bacteria and viruses – see future methods. A current EnviroLink Tools project is comparing standard bioaccumulation models for viruses with new models that explicitly account for uptake and depuration (McBride 2016).
Computational approaches and metrics derived
<u>National standards / guidelines and consistency:</u> Limited to the standardised thresholds set out below. **<u>Opportunities:</u>** Possible to incorporate a Quantitative Microbial Risk Assessment approach (QMRA) to developing metrics. Current QMRA modelling for contamination of shellfish potentially affected by discharges of treated wastewater uses a bioaccumulation approach, ignoring uptake and depuration. A model that does take explicit cognisance of the processes has been developed and is currently being compared with the former approach (G McBride, NIWA, Envirolink 'Tools' project, MBIE contract C10X1610). **<u>Caveats and recommendations</u>**: Natural patchiness in the distribution of faecal indicator bacteria can impede the ability to identify trends over time. This may be less of an issue for shellfish samples than for water samples, since the filter-feeding shellfish can be good integrators of patchiness in the water.
Thresholds
The Ministry for the Environment (2003) Guidelines for recreational shellfish gathering include for example, median concentration of faecal coliforms taken over a shellfish gathering season shall not exceed a most probable number (MPN) of 14 per 100 mL and not more than 10% of samples should exceed an MPN of 43 per 100 mL.

The microbiological limits outlined in the New Zealand Food Safety Authority Animal Products (Specifications for Bivalve Molluscan Shellfish) Notice 2006 for commercial shellfish quality state the *E.coli* median MPN of shellfish samples must not exceed 230 *E.coli* per 100 g and not more than 10 percent of the samples must exceed an MPN of 700 per 100 g. The limit can vary depending on proximity to point sources of contamination.
Emerging and prospective future methods
As summarised in Green and Cornelisen (2016), emerging technologies for monitoring FIB may replace or complement culture based tests for FIB as they become validated. Tests for FIB that are faster than the current culture based tests will address the current challenge around delayed results; typically results using standard culture methods cannot be produced for at least 24 hours following sample collection. As is the case for water samples, Microbial Source Tracing (MST) using DNA-based markers can also be applied in shellfish. There have been a number of studies trialing methods of extraction and evaluating marker detection in shellfish for a range of molecular markers, including source-specific bacteria and viruses (e.g., Kirs and Cornelisen 2011). Such work is now linked to research in the NZ Safe Seafood programme that involves virus detection in shellfish and development of models for forecasting transport and fate of upstream faecal contaminants in shellfish harvest areas.
References
Bolton-Ritchie L, Greenfield S., Madarasz-Smith, A., Milne, J., Stevenson M. and Walker J. (2013) Recreational water quality: A practitioner's discussion on the limitations of the 2003 national guidelines. Unpublished paper prepared on behalf of the regional sector's Coastal Special Interest Group (C-SIG) and the Surface Water Integrated Management (SWIM) Group. Cornelisen, C., Kirs, M., Gilpin, B.E., Scholes, P.E. (2012) Microbial Source Tracking (MST) Tools for Water Quality Monitoring. *Cawthron client report prepared for Regional Councils and the Coastal Special Interest Group*, No. 2047: 28. Dudley, B., Zeldis, J., Burge. O. (2017) New Zealand Coastal Water Quality Assessment. Prepared for Ministry for the Environment. NIWA CLIENT REPORT No: 2016093CH. Green, M.O., Cornelisen, C.D. (2016) Marine Water Quality Standards for the Waikato Region – Literature Review. Report WRC1507–1, Streamlined Environmental, Hamilton, 123 pp. Kirs, M., Cornelisen, C. (2011) MST Technology for Shellfish; Final Report. *Prepared for Seafood Innovations Ltd*. Cawthron Report No. 2002. 36 p. McBride, G. (2016) Marine contact recreation water and shellfish guidelines Envirolink Tools proposal 2016/17. Prepared for the Ministry for Business, Innovation and Excellence on behalf of the regional sector's Coastal Special Interest Group. Contract C01X1610. Milne, J., Madarasz-Smith, A., Davie, T. (2017) Recreational water quality monitoring and reporting; A position paper prepared for the NZ regional sector. 28 p. Ministry for the Environment (2003) Microbiological Water Quality Guidelines for Marine and Freshwater Recreational Areas. Ministry for the Environment, Wellington, New Zealand. MPI 2013; Enumeration of Escherichia coli in Bivalve Molluscan Shellfish MPI Method Version 9.

Emerging contaminants (or Contaminants of Emerging Concern)

Definition: Any synthetic or naturally occurring chemical (or any microorganic) that is not commonly monitored in the environment but has the potential to enter the environment and cause known or suspected adverse ecological and/or human health effects.

There is increasing concern about so-called 'contaminants of emerging concern', including many 'micropollutants' — small, persistent and biologically active substances — among them certain pesticides, industrial chemicals, pharmaceuticals and personal care products. The contamination of environmental compartments — such as surface water, groundwater and soil — with these chemicals can have adverse effects on aquatic organisms, and on human health if they accumulate in seafood or get into drinking water.

Under the European Union (EU) mandated Water Framework Directive (WFD), environmental quality standards (EQS) have been established for 45 so-called 'priority substances' and eight other pollutants. When the Directive on Environmental Quality Standards was amended in 2013, a watch list mechanism was established to require temporary monitoring of other substances for which evidence suggested a possible risk to or via the environment, to inform the selection of additional priority substances. In addition, the 2013 Directive identified three substances (the natural hormone oestradiol (E2) and two pharmaceuticals — the anti-inflammatory diclofenac and the synthetic hormone ethinyl oestradiol (EE2), used in contraceptives) for inclusion in the first watch list to facilitate the determination of appropriate measures to address the risk posed by those substances (Barbosa et al. 2016).

The first watch list was adopted in 2015 (in Decision 2015/495) and also includes the following chemicals:

- the natural hormone oestrone (E1)
- three (macrolide) antibiotics
- several pesticides
- a UV filter (a chemical that prevents UV light getting though, as used in sun cream), and
- an antioxidant used as a food additive.

Of particular interest are also: (i) engineered nanomaterials; and (ii) pharmaceuticals and personal care products; which have increasing usage in society.

The process to establish a short-list of contaminants of emerging concern should include: a screening level risk assessment of pollution sources; the types and potential loadings of chemical contaminants; ecological and human health toxicity of chemicals; and likely fate pathways in the receiving environment. In New Zealand, such risk assessments have rarely been undertaken because of the lack of information on components of the assessment process. Historical assessments have been undertaken for pesticides in horticultural in surface waters (Wilcock 1989; Wilcock and Close 1990; Holland and Rahman 1999) and groundwater environments (Close 1993; Close 1996; Close and Flintoft 2004) – though more recent comprehensive pesticide-related risk assessments are lacking. Ahrens (2008) undertook a comprehensive review of organic chemicals of potential environmental concern in use in Auckland – which include toxicological hazard ranking of the various contaminant classes.

Elevated concentrations of other "traditional" and emerging contaminants found in sediments (e.g., metals and metalloids such as mercury and arsenic, and PAHs (McHugh and Reed 2006); pharmaceuticals and emerging contaminants (Stewart 2013; Stewart et al. 2014; Stewart et al. 2016); are generally expected to be transported to the estuary as sediment-associated contaminants rather than as elevated dissolved concentrations – which could exceed their respective water quality guidelines. As such, controls to manage sediment-associated contaminants in stormwaters and sewage overflows will be expected to result in reduced concentrations of these contaminants entering the marine environment. Other organic and organo-metallic formulations enter estuaries from activities occurring within the estuary, such as antifouling agents and co-biocides in antifouling products (Boxall et al. 2000; Stewart 2003; Stewart 2006; Stewart and Conwell 2008; Stewart et al. 2008; Gadd et al. 2011). More general reviews are available which provide an assessment of limits and guidelines available for classifying New Zealand estuaries and coastal waters (Green and Cornelisen 2016; Williamson et al. 2017).

A new group of chemicals are emerging throughout the world as being of potential environmental concern, based on their toxicity, persistence, and widespread or on-going use. These have been termed Chemicals of Potential Environmental Concern (CPEC) or Emerging Chemicals of Concern (ECC). In contrast to the "priority pollutants", many CPECs have a lower environmental hazard profile. Notably, many CPECs have lower acute toxicity than Priority Pollutants (PP). Nevertheless, some CPECs have the potential to exert chronic effects by being neuroactive or acting as hormone mimics (endocrine disrupting chemicals). Some are associated with high production volumes, so there is a potential for accumulation of these chemicals in estuarine receiving environments, with unknown consequences, with risks elevated in intensively urbanized estuaries. The differences between PP and CPEC are summarised in Table 1.

Table 1 - Comparison of risk profile of priority pollutants and emerging chemicals of potential environmental concern (Williamson et al. (2017) adapted from Ahrens (2008)).

Property	Priority Pollutants (PP)	Chemicals of Potential Environmental Concern (CPEC)
Toxic effects and mode of action	Acute and chronic	Most not likely to be acutely toxic at environmental doses, but potentially bioactive (e.g., estrogenic, neuro-active), sometimes at very low concentrations
Environmental concentrations	Frequently monitored; stable or decreasing (except Zn, Cu, PAH in urban stormwater)	Not frequently monitored, assumed to be increasing
Persistence	High	Variable: unknown, low, medium, high
Bioaccumulation potential	High	Variable: unknown, low, medium, high
Sources	Mainly industrial and agricultural; building materials and vehicles; few domestic (i.e., sewage)	Some industry and agriculture runoff; mostly domestic (via sewer overflows, wastewater discharges)
Existing water quality guideline	Yes	No

Discharge regulated	Often (but not in diffuse runoff in NZ)	Rarely
Detection and quantification	Relatively easy; methods are well established	Often difficult and expensive to measure; focus often on use of biomarker techniques
Examples	As, Cd, Hg, Pb, DDT, PCB, PAH, dioxin, Cu, Zn	Surfactants, plasticizers, disinfectants, modern pesticides, flame retardants, hormones, cosmetics, new antifouling paints, medicines, veterinary medicines

A comprehensive list of priority substances is provided in the EU Water Framework Directive (WFD) Strategy on Priority Substances (Directive 2000/60/EC; replaced by Annex II of the Directive on Environmental Quality Standard (Directive 2008/105/EC)[19]) (Table 2). These include seven substance classes which form the basis of EU environmental monitoring programmes for chemical contaminants (European Commission 2010; European Commission 2014). The ANZECC (2000) guidelines provide marine water guidelines for metals and metalloids, aromatic hydrocarbons (including PAHs) and a limited range of organic contaminants. The ANZECC (2000) marine water quality trigger values for some of the EU priority substances are shown in Table 3, with many being "low reliability" reflecting the lack of data for the original derivation process. The guidelines are current being revised and updated (Warne et al. 2014; Warne et al. 2015), with that process targeting several marine priority substances.

Table 2: List of priority substances in the European Union water policy (Directive 2008/105/EC).

Group	Chemicals
Metals and metalloids	Cadmium, lead, mercury, nickel, tributyltin, and their compounds
Aromatic hydrocarbons (including polycyclic aromatic hydrocarbons (PAHs))	Anthracene, benzene, fluoranthene, naphthalene, benzo(a)pyrene, benzo(b)fluoranthene, benzo(g,h,i)perylene, benzo(k)fluoranthene, indeno(1,2,3-cd)pyrene, trichlorobenzene, pentachlorobenzene
Pesticides (insecticides, herbicides, fungicides)	Alachlor, atrazine, chlorpyrifos, chlorfenvinphos, 1,2-dichloroethane, dichloromethane, diuron, endosulfan, hexachlorobutadiene, hexachlorobenzene, hexachlorocyclohexane, isoproturon, pentachlorophenol, simazine, trifluralin
Flame retardants	Brominated diphenylether (pentabromodiphenylether congeners 28, 47, 99, 100, 153, 154)
Chloroalkanes	C10-13, trichloromethane (chloroform)
Alkylphenols	Nonylphenol, octylphenol
Plasticizer	Di(2-ethylhexyl)phthalate (DEHP)

[19] http://ec.europa.eu/environment/water/water-dangersub/pri_substances.htm; http://ec.europa.eu/environment/water/water-framework/priority_substances.htm

Table 3: Marine water quality trigger values (95% level of species protection) for organic priority substances from ANZECC (2000) (from Green and Cornelisen (2016)).

Chemical	Trigger value (µg/L), freshwater	Trigger value (µg/L), marine
nonylphenols	0.1 [A]	1.0 [B]
dimethylphthalate	3700 [C]	3700 [B]
diethylphthlate	1000 [A]	900 [D]
di-n-butyl phthalate	35 [A]	25 [B]
di-2-(ethylhexyl)phthalate	1 [A]	1 [B]
chlorpyrifos	0.01 [E]	0.009 [F]
diuron	0.20 [A]	1.8 [B]
glyphosate	1200 [C]	370 [B]
Linear alkylbenzene sulfonates (LAS)	280 [E]	0.1 [B]
Alcohol ethoxylated sulphates (AES)	650 [E]	650 [B]
Alcohol ethoxylate surfactants (AE)	140 [E]	140 [B]

[A] Low reliability trigger value for fresh water (µg/L), indicative interim working level only.
[B] Low reliability trigger value for marine water (µg/L), indicative interim working level only.
[C] Moderate reliability trigger value for fresh water (µg/L).
[D] Moderate reliability trigger value for marine water (µg/L), indicative interim working level only.
[E] High reliability trigger value for fresh water (µg/L).
[F] High reliability trigger value for marine water (µg/L).

Three reports have reviewed the literature on emerging organic contaminants (EOCs) and their relevance to New Zealand estuarine environments (Ahrens 2008; Tremblay et al. 2011; Stewart et al. 2016). The recent report prepared for Auckland Council, Environment Canterbury and Greater Wellington Regional Council (Stewart et al. 2016) summarises information required by regional councils in New Zealand to address concerns around the environmental risks for adverse effects of EOCs. The report also recommends an approach for councils to target monitoring efforts to a tiered suite of (primarily sediment) indicators of EOCs.

There is currently no national strategy in New Zealand for managing EOCs. Internationally, regulatory bodies around the world are starting to impose restrictions or bans on selected EOCs, with many more being placed on watch lists for future assessments. Some BDE[20] and PFOS[21]/PFOA[22] have been identified for elimination or reduction by their inclusion in the Stockholm Convention on Persistent Organic Pollutants. Within New Zealand, the Environmental Protection Authority (EPA) has the ability to re-assess approvals for EOCs, and recently revoked approvals for the antifouling co-biocides irgarol and chlorothalonil, and 18 veterinary medicine and insecticide products, including carbaryl, chlorpyrifos and diazinon (Stewart et al. 2016).

New Zealand EOCs research relevant to estuaries includes a literature review, which included hazard risk categories and recommended monitoring for urban-sourced EOCs (Ahrens 2008). This was followed up by field analysis of EOCs around the Auckland marine environment (Stewart et al. 2008). Archived sediments sourced from the Auckland urban study were analysed for a suite of 46 pharmaceuticals (Stewart 2013; Stewart et al. 2014). Passive sampling devices (PSDs) for heavy metals and EOCs have also been evaluated as a potential replacement for shellfish biomonitoring programmes (Stewart et al. 2016).

[20] Brominate diphenylethers
[21] Perfluorooctanesulfonic acid
[22] Perfluorooctanoic acid

Developing emerging contaminant attributes is challenging because of the wide range of chemicals of potential concern – many of which are present at trace level concentrations – and complicated by the time and spatial (within and between) variability of contaminant exposure conditions in estuaries (i.e., internal vs external sources) and the suitability of local reference (benchmarking) sites. Guidelines for ECs in water are limited and rare for sediments – which potentially requiring benchmarking to local reference sites as an arbitrary exposure measure. The nature of the EC hazards will differ significantly between estuaries in different regions/catchments – requiring the suite of ECs monitored to be site-specific to usefully link to potential ecosystem or human health concerns.

Sampling design

National standards / guidelines and consistency: Monitoring for ECs can be undertaken using chemical monitoring of waters, sediments or biota. The choice of the approach will be dependent on the objectives of the monitoring programme. The use of shellfish or fish provides direct measure of bioavailable contaminant (addressed in the Shellfish contaminants factsheet) – but is potentially limited to chemical compounds which significantly bioaccumulate and are not metabolised and excreted by the organism's detoxification processes. This factsheet addresses chemical monitoring approaches for water and sediments in the estuarine and coastal environment.

Chemicals of Potential Environmental Concern (CEPCs)
Ahrens (2008) conducted a very comprehensive review of CPEC that are emerging in the world's literature. Based on this review, CPEC do not appear to reach environmental concentrations able to exert acute toxicity effects on biota. However, if moderately elevated concentrations are present, or bioavailability is enhanced with long-term exposure, there is the possibility of chronic effects on organism health. Because they are likely to occur in mixtures, there is the possibility of additivity of toxicity of chemicals with a common mode of action, such as endocrine disrupting compounds (EDCs). Thus, while the environmental concentrations may fall below the levels where a specific chemical is known to affect organisms, these chemicals may act in concert, producing an additive or synergistic adverse effects.

In addition to urban stormwater as a potential source for such CPECs as pesticides, plasticizers, and petroleum products, Ahrens (2008) identified many other potential sources in the urban landscape including marinas, sewage outfalls, combined sewage overflows, landfill leachate, and agricultural runoff (Table 4).

CPECs have been surveyed in Auckland on two occasions, and these surveys characterise typical concentrations and distributions (Stewart et al. 2008; Stewart 2013; Stewart et al. 2014; Stewart et al. 2016). In addition, endocrine disrupting chemical (EDC) measurement and assessment have been reviewed in relation to Auckland (Singhal et al. 2009).

A recent report by Stewart et al. (2016) reviews emerging organic contaminants (EOCs) relevant to New Zealand's estuarine environments and recommends a tiered suite of (primarily sediment) indicators of EOCs. The multiple major sources and pathways of EOCs to the marine environment are illustrated in Figure 1 and summarised in Table 4. It should be noted that this does not address other classes of ECs, such as nanomaterials, which may pose risks to estuarine and coastal ecosystems (Klaine et al. 2008). The "core" suite of EOCs recommended by Stewart et al. (2016) as a Tier I assessment of sediments is shown in Table 5.

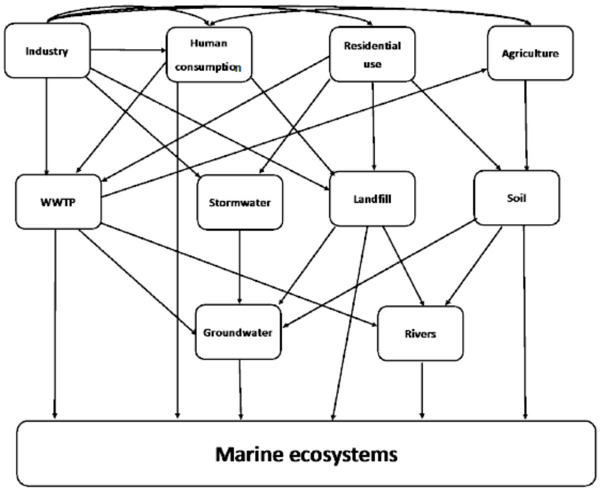


Figure 1: Sources and pathways of EOCs into the marine receiving environment (from Stewart et al. (2016)).

Table 4: Classes of EOCs by major sources (from Stewart et al. (2016)).

EOC Class	Sewage	Stormwater	Landfill leachate	Agriculture and Horticulture	Aquaculture /Marine industry	Recreation
Pharmaceuticals	☐	☐	☐			☐
Plasticisers	☐	☐	☐			
Antimicrobials	☐	☐	☐			
Corrosion inhibitors	☐	☐				
Flame retardants	☐	☐	☐			
Surfactants	☐	☐	☐	☐		
UV-filters	☐	☐	☐	☐		☐
Steroid hormones	☐			☐		
Musk fragrances	☐	☐	☐			
PFOAs etc	☐	☐	☐			
Veterinary medicines				☐	☐	
Pesticides				☐		
Antifouling cobiocides					☐	

Table 5: "Core" list of "marker" EOCs recommended for initial phase (Tier 1) of sediment monitoring (from Stewart et al. (2016)).

Class	Representative EOC[a,b]	CAS	Major Sources[c]	Reason[d]
Flame retardants	BDE47	5436-43-1	SEW,SW,LF	1,2,3,5,6
	BDE99	60348-60-9	SEW,SW,LF	1,2,3,5,6
	BDE209	1163-19-5	SEW,SW,LF	1,2,3,5,6
	TDCP	13674-87-8	SEW,SW,LF	1,2,4,6
	TPP	115-86-6	SEW,SW,LF	1,2,4,6
	TCPP	13674-84-5	SEW,SW,LF	1,2,4,6
Plasticisers	DEHP	117-81-7	SEW,SW,LF	2,3,5
	BBP	85-68-7	SEW,SW,LF	2,3,5
	Bisphenol A	80-05-7	SEW,SW,LF	1,5
Surfactants	Nonylphenol	84852-15-3	SEW,SW,LF,AG	1,2,3,5,6
	LAS	25155-30-0	SEW,SW,LF,AG	2,4
Perfluorinated compounds	PFOS/PFOA	1763-23-1/335-67-1	SEW,SW,LF	1,2,4,6
Musk fragrances	Galaxolide	1222-05-5	SEW,SW,LF	2,3,4,6
	Tonalide	21145-77-7	SEW,SW,LF	2,3,4,6
Pesticides	Glyphosate/AMPA	1071-83-6	AG	1,2,3,5
Neonicotinoid insecticide	*Imidacloprid*	138261-41-3	AG	1,4
Pyrethroid insecticide	Bifenthrin / Permethrin	82657-04-3	SEW,SW,LF,AG	2,4
Pharmaceuticals	Acetaminophen	103-90-2	SEW,SW,LF,REC	2,3,5
	Diclofenac	15307-86-5	SEW,SW,LF,REC	2,3,5
	Ibuprofen	15687-27-1	SEW,SW,LF,REC	2,5
	Carbamazepine	298-46-4	SEW,SW,LF,REC	2,4
Steroid estrogen	Estrone	53-16-7	SEW,AG	4,5
Personal Care Products	Triclosan	3380-34-5	SEW,SW,LF	1,2,6
	Methyltriclosan	1/01/1940	SEW,SW,LF	1,2,5,6
Preservative	Methylparaben	99-76-73	SEW,SW,LF	2,5
Corrosion inhibitor	*Benzotriazole*	95-14-7	SEW,SW	2,4

[a] BDE = brominated diphenyl ether; DEHP = Bis(2-ethylhexyl)phthalate; BBP = benzyl butyl phthalate; LAS = linear alkylbenzene sulfonate; PFOS = perfluorooctanesulfonic acid; PFOA = perfluorooctanoic acid; TCPP = Tris (1-chloro-2-propyl) phosphate; TDCP = Tris[2-chloro-1-(chloromethyl)ethyl]phosphate; TPP = Triphenylphosphate.

[b] Currently no laboratory capability for analysis of italicized EOCs in New Zealand.

[c] Major sources see Table 3. SEW = sewage; SW = stormwater; LF = landfill; AG = agriculture/horticulture; AQ = aquaculture; REC = recreation.

[d] 1 Initiative to remove. Stockholm Convention (POPs) or individual initiatives; 2 High production chemical; 3 Highest concentrations detected in urban marine receiving environment; 4 Knowledge gap (not previously monitored); 5 Previously detected in NZ marine sediments; 6 Persistent Bioaccumulative and Toxic (PBT).

Spatial design

Optimising a monitoring programme for land-use related effects would need to consider the suite of hazards likely to result in EOCs entering into a specific estuarine environment in order to potentially refine the "core" list of Tier 1 EOCs (Table 5).

Additionally, site-specific consideration will be required to determine sediment monitoring sites within an estuary. Ideally, the use a hydrodynamic model of an estuary would assist in characterising the location of monitoring sites in relation to contaminant sources within and external to the estuary and for local reference sites (e.g., Xu et al. (2018)). This would aid in establishing anticipated contaminant exposure gradients and reduce the potential for redundancy in site selection.

Potential bottlenecks: C monitoring data is limited in terms of spatial extent and number of replicates, selection of surveillance and reference sampling sites. If a long-term monitoring programme is implemented (e.g., 5-yearly), then a long period will be required before statistical data (e.g., median, 95th percentile) can be collected for between site comparisons, or for statistical detection of trends.

Opportunities: Depositional fine sediments are potentially good integrators of chemical contaminants within an estuarine environment. Combining programmes for sediment monitoring for traditional chemical contaminants, ECs and shellfish biomonitoring programmes for chemical contaminants and faecal indicator bacteria (FIB) in shellfish to ensure common sites and concurrent sampling will provide greatest information. Utilising exposure gradients within the estuary will provide an ability to link contaminants with sources. Estuarine monitoring programmes will also require integration with a catchment monitoring programmes in order to distinguish legacy contaminant present within the estuarine sediments, local within estuary resuspension and transport and catchment/landuse loads.

Combined programmes will provide greater information for ecological and human health risk assessment.

Caveats and recommendations: Site selection for surveillance and reference monitoring sites are critical to the success of a monitoring programme. Specific-site locations and knowledge of estuary hydrodynamics are required to link shellfish contaminants with local catchment inputs.

Analytical method detection limits (MDLs) may be limited for some EC, with consistent methodologies and MDLs being required throughout the monitoring programme to provide a robust ability for detection of environmental exposures. Specific high sensitivity methods may be applicable to specific contaminants of concern which present a high potential risk to the environment.

Sampling procedures

National standards / guidelines and consistency: No national standards. Standard methods for sediment sampling are available (Hickey et al. 1995; Mills and Williamson 2008; Simpson and Batley 2016). Compositing of sediment cores from a monitoring site may be advisable to obtain a single screening concentration for each of the suite of contaminants to reduce the cost of the monitoring programme. If elevated EOC concentrations are detected then replicated sediment samples from archived sediments can be subsequently analysed to obtain a measure of variability for the site. No national standards for EC thresholds are available for determining adverse effects in sediments.

Techniques are being developed for the use of passive samplers for application to waters and sediments (e.g., using DGT (diffusive gels in thin film) techniques, Stewart et al. (2016)). Quantitative measurements from DGTs for water and sediment pore waters can be compared with the limited range of guidelines available to determine likely adverse effects. Guidance is available for application of passive samplers for monitoring chemical contaminants in sediments (Burgess 2012; Ghosh et al. 2014).

Spatial design of monitoring programme and criteria for reference sites requires standardization.

Potential bottlenecks: Bottlenecks include those as outlined under sampling design and relate to spatial and temporal replication of samples.

Opportunities: Monitoring programmes can include sediment analyses for traditional contaminants and ECs; potential to link with shellfish and/or fish biomonitoring programmes.

Caveats and recommendations: A standardised "core" range of ECs and comparable MDLs will improve quality control and inter-comparison between estuaries in different regions.

Laboratory analyses

National standards / guidelines and consistency: No national standards. Chemical analyses for tissue and shell are available for a wide range of common chemical contaminants from commercial laboratories (e.g., Hill Laboratories Ltd, AssureQuality, Northcott Consulting Limited). Specialist analytical laboratories are generally required for trace organic compounds with environmentally relevant MDLs (Stewart et al. 2016).

Standard sediments are available to include in the analytical QC procedures for some ECs. Spiked sediment matrices may need to be specifically prepared for most of the EOCs.

Potential bottlenecks: Chemical contaminant method detection limits may differ between analytical laboratories. ECs generally not routinely analysed in standard suite of chemical analyses – requiring specialist laboratory services or development of new techniques. Sediments can be archived for subsequent analyses.

Opportunities: Chemical contaminant monitoring in sediments can be combined with shellfish biomonitoring contaminants. However, the time scales of the two monitoring programmes will differ – with shellfish requiring longer term exposure for chemical contaminant bioaccumulation. **Caveats and recommendations**: Robust QA procedures should be incorporated into the chemical analytical procedures. Archived samples should be retained from key sites for potential analysis for additional 'new' contaminants once suitable analytical methods are available, or for reanalysis of existing contaminants using updated methods with improved detection limits. Cost for analysis will be high for many of the EOCs in the "core" Tier I list (Stewart et al. 2016) (indicatively $2800/sample for multi-sample batches, G. Olsen, NIWA, pers com).
Computational approaches and metrics derived
National standards / guidelines and consistency: No sediment quality guidelines available for EOCs. Comparative data from reference sites can be used for EOCs to establish effects bands – which would be arbitrary and not related to thresholds for adverse effects. Some water quality guidelines are available for EOCs (existing and updated ANZECC (2000)). **Potential bottlenecks:** Metrics derived based on comparative data with reference sites will be arbitrary. Comparisons between sites and over time may be limited by analytical detection limits – which may differ between laboratories and over time as methods change. Analytical detection limits for key contaminants of concern need to be sufficiently low to detect chemical exposures at concentrations which are environmentally relevant to potential adverse effects on aquatic organisms. Need statistical and reporting methods that deal with data measurements less than the detection limits. Multiple estuarine monitoring sites and a spatial component of the monitoring design will be required for establishment of local reference sites and distinguishing catchment from local legacy and harbour-generated sources. **Opportunities:** Multiple metrics of exposure and 'health' effects on organisms can be used if EC sediment monitoring combined with traditional sediment contaminants and biomonitoring for tissue body-burdens and toxicity testing for adverse effects. **Caveats and recommendations**: Standard monitoring programmes for trend detection should be undertaken at comparable seasons to standardize time-varying catchment loads. The suite of EOC contaminants should be refined in different estuaries based on the likely chemical hazards present in the estuary catchments. Compositing of replicated sediment samples from individual sites should be used for initial screening to reduced monitoring programme costs.
Thresholds
Ecological Marine waters: ANZECC (2000) and updates presently being derived; various international jurisdictions. Marine sediments: None. *Food safety* Food safety assessments can only be made on tissue from fish or shellfish. The Australia New Zealand Food Standards Code (Food Standards Australia and New Zealand 2015) prescribes maximum levels for arsenic (As), cadmium (Cd), lead (Pb), mercury (Hg), polychlorinated biphenyls (PCB), histamine and marine biotoxins in seafoods. Standard 1.4.1, Contaminants and natural toxicants, sets out the maximum levels (MLs) of specified metal and non-metal contaminants and natural toxicants in nominated foods (http://www.foodstandards.gov.au/foodstandards/foodstandardscode.cfm).

There are no food safety standards for the "core" EOCs (Table 5) or nanomaterials (Klaine et al. 2008).
Emerging and prospective future methods
Developing emerging contaminant attributes is challenging because of the wide range of chemicals of potential concern – many of which are present at trace level concentrations – and complicated by the time and spatial (within and between) variability of contaminant exposure conditions in estuaries (i.e., internal vs external sources) and the suitability of local reference (benchmarking) sites. Guidelines for ECs in water are limited and rare for sediments – potentially requiring benchmarking to local reference sites as an arbitrary exposure measure. The nature of the EC hazards will differ significantly between estuaries in different regions/catchments – requiring the suite of ECs monitored to be site-specific to usefully link to potential ecosystem or human health concerns. The development of passive sampling techniques for waters and sediments will go some way to addressing time-varying exposures, for targeted "high-risk" contaminants, and to provide a measure of "bioavailable" dissolved fractions. The use of passive devices, such as DGT samplers, will provide a practical way for surveillance monitoring by management agencies. For many ECs, a primary objective of the monitoring will be to robustly demonstrate an absence of those specific contaminants, or contaminant classes, in the estuarine environment. Therefore, reporting and classification systems must accommodate the negative results showing an absence of those ECs. The development of an attribute classification based on ECs will be challenging – with major challenges to provide a nationally applicable system of standards. Because most of the ECs are likely to be present at concentrations below known adverse effects thresholds in sediments there is no basis to expect an effects-based classification system for ecosystem "health" protection. Therefore, at present, the monitoring for ECs in estuarine environments is probably best suited to surveillance monitoring with the programme design targeting "hot spots" based on known or anticipated catchment or internal contaminant loads.
References
Ahrens, M. (2008). Review of organic chemicals of potential environmental concern in use in Auckland. No. HAM2007-141. NIWA client report prepared for Auckland Regional Council, pp. 180. ANZECC (2000). Australian and New Zealand guidelines for fresh and marine water quality. National Water Quality Management Strategy Paper No. 4, Australian and New Zealand Environment and Conservation Council and Agriculture and Resource Management Council of Australia and New Zealand, Canberra, Australia. Barbosa, M.O.; Moreira, N.F.F.; Ribeiro, A.R.; Pereira, M.F.R.; Silva, A.M.T. (2016). Occurrence and removal of organic micropollutants: An overview of the watch list of EU Decision 2015/495. *Water Research 94*: 257-279. Boxall, A.B.A.; Comber, S.D.; Conrad, A.U.; Howcroft, J.; Zaman, N. (2000). Inputs, monitoring and fate modelling of antifouling biocides in UK estuaries. *Marine Pollution Bulletin 40*: 898-905. Burgess, R.M. (2012). Guidelines for Using Passive Samplers to Monitor Organic Contaminants at Superfund Sediment Sites. No. EPA/600/R-11/115. U.S. Environmental Protection Agency, Washington, DC. pp. 32. Close, M.E. (1993). Assessment of pesticide contamination of groundwater in New Zealand. 1. Ranking of regions for potential contamination. *New Zealand Journal of Marine and Freshwater Research 27*: 257-266. Close, M.E. (1996). Survey of pesticides in New Zealand groundwaters, 1994. *New Zealand Journal of Marine and Freshwater Research 30*: 455-461. Close, M.E.; Flintoft, M.J. (2004). National survey of pesticides in groundwater in New Zealand - 2002. *New Zealand Journal of Marine and Freshwater Research 38*: 289-299. European Commission (2010). WFD-CIS Guidance Document No. 25 Guidance on chemical monitoring of sediment and biota under the Water Framework Directive. No. Technical report: 2010-041. Office for Official Publications of the European Communities, Luxembourg. pp. 74.

European Commission (2014). WFD-CIS Guidance Document No. 32 Guidance on biota monitoring (the implementation of EQS_{BIOTA}) under the Water Framework Directive. No. Technical report: 2014-083. Office for Official Publications of the European Communities, Luxembourg. pp. 87.

Food Standards Australia and New Zealand (2015). Australia, New Zealand food standards code - Standard 1.4.1 - Contaminants and Natural Toxicants (http://www.foodstandards.gov.au/code/Pages/default.aspx). www.foodstandards.gov.au.

Gadd, J.; Depree, C.; Hickey, C.W. (2011). Relevance to New Zealand of the OECD Emission Scenario Document for Antifouling Products: Phase 2 Report. No. ERM10201; HAM 2011-005. NIWA report for the Environmental Protection Authority (EPA), Wellington. pp. 108.

Ghosh, U.; Kane Driscoll, S.; Burgess, R.M.; Jonker, M.T.O.; Reible, D.; Gobas, F.; Choi, Y.; Apitz, S.E.; Maruya, K.A.; Gala, W.R.; Mortimer, M.; Beegan, C. (2014). Passive sampling methods for contaminated sediments: Practical guidance for selection, calibration, and implementation. *Integrated Environmental Assessment and Management 10*: 210-223.

Green, M.O.; Cornelisen, C.D. (2016). Marine Water Quality Standards for the Waikaoto Region - Literature Review. No. WRC1507-1. Streamlined Environmental report prepared for Waikato Regional Council, Hamilton. pp. 123.

Hickey, C.W.; Roper, D.S.; Holland, P.T.; Trower, T.M. (1995). Accumulation of organic contaminants in two sediment dwelling shellfish with contrasting feeding modes: deposit - (*Macomona liliana*) and filter-feeding (*Austrovenus stutchburyi*). *Archives of Environmental Contamination and Toxicology 29*: 221-231.

Holland, P.; Rahman, A. (1999). Review of Trends in Agricultural Pesticide Use in New Zealand. No. MAF Policy Technical Paper 99/11. pp.

Klaine, S.J.; Alvarez, P.J.J.; Batley, G.E.; Fernandes, T.F.; Handy, R.D.; Lyon, D.Y.; Mahendra, S.; McLaughlin, M.J.; Lead, J.R. (2008). Nanomaterials in the environment: Behavior, fate, bioavailability, and effects. *Environmental Toxicology and Chemistry 27*: 1825-1851.

McHugh, M.; Reed, J. (2006). Marine Sediment Monitoring Programme - 2005 Results. No. ARC 06109: TP316. NIWA report for Auckland Regional Council. Auckland Regional Council Technical Publication No. 316, Auckland. pp. 197.

Mills, G.N.; Williamson, R.B. (2008). The Impacts of Urban Stormwater in Auckland's Aquatic Receiving Environment A Review of Information 1995 to 2005. No. Auckland Council technical report TR 2008/029. Prepared by Diffuse Sources Ltd and Geosyntec Consultants for Auckland Regional Council, Auckland. pp. 241.

Simpson, S.L.; Batley, G.E., (Eds.) (2016). *Sediment quality assessment : a practical guide*. CSIRO Publishing, Clayton, Vic.

Singhal, N.; Song, Y.; Johnson, A.; Swift, S. (2009). Estrogenic Endocrine Disrupting Compounds. No. Auckland Regional Council Technical Report Number TR 2010/005. Prepared by UniServices for Auckland Regional Council, Auckland. pp.

Stewart, C. (2003). Antifouling co-biocides in New Zealand coastal waters. . No. Report prepared for Ministry for the Environment, Wellington. pp.

Stewart, C. (2006). Antifouling co-biocides in New Zealand coastal waters: 2006 resurvey. No. Prepared for Ministry for the Environment, Wellington. pp. 49.

Stewart, C.; Conwell, C. (2008). Temporal changes in contaminated sediment: A neglected aspect of ecological risk assessment? *Integrated Environmental Assessment and Management 4*: 517-519.

Stewart, M. (2013). Pharmaceutical residues in the Auckland estuarine environment. No. Auckland Council technical report, TR2013/002. Prepared by NIWA for Auckland Council. , pp. 51.

Stewart, M.; Ahrens, M.; Olsen, G. (2008). Analysis of chemicals of emerging environmental concern in Auckland aquatic sediment. ARC08226. Hamilton, NIWA. *HAM2008-057*: various pagings

Stewart, M.; Cameron, M.; McMurtry, M.; Sander, S.G.; Benedict, B.; Graham, L.; Hosie, M.; Green, T. (2016). Development of passive sampling devices for bioavailable contaminants of current and emerging concern: Waitemata Harbour case study. *New Zealand Journal of Marine and Freshwater Research 50*: 526-548.

Stewart, M.; Northcott, G.L.; Gaw, S.; Tremblay, L.A. (2016). An Update on Emerging Organic Contaminants of Concern for New Zealand with Guidance on Monitoring Approaches for Councils. No. Technical Report 2016/006. Prepared by Streamlined Environmental Ltd, Northcott Research Consultants Ltd, University of Canterbury, Cawthron Institute and the University of Auckland for Auckland Council, Greater Wellington Regional Council and Environment Canterbury Regional Council, Auckland. pp. 120.

Stewart, M.; Olsen, G.; Hickey, C.W.; Ferreira, B.; Jelić, A.; Petrović, M.; Barcelo, D. (2014). A survey of emerging contaminants in the estuarine receiving environment around Auckland, New Zealand. *Science of the Total Environment 468–469*: 202-210.

Tremblay, L.A.; Stewart, M.; Peake, B.M.; Gadd, J.B.; Northcott, G.L. (2011). Review of the Risks of Emerging Organic Contaminants and Potential Impacts to Hawke's Bay. No. 1973. Cawthron report prepared for Hawke's Bay Regional Council, Nelson. pp. 39.

Warne, M.S.; Batley, G.E.; Braga, O.; Chapman, J.C.; Fox, D.R.; Hickey, C.W.; Stauber, J.L.; Van Dam, R. (2014). Revisions to the derivation of the Australian and New Zealand guidelines for toxicants in fresh and marine waters. *Environmental Science and Pollution Research 21*: 51-60.

Warne, M.S.; Batley, G.E.; van Dam, R.A.; Chapman, J.C.; Fox, D.R.; Hickey, C.W.; Stauber, J.L. (2015). Revised Method for Deriving Australian and New Zealand Water Quality Guideline Values for Toxicants. Updated May 2017. No. Prepared for the Council of Australian Government's Standing Council on Environment and Water (SCEW) Department of Science, Information Technology, Innovation and the Arts, Brisbane, Queensland. pp. 46.

Wilcock, R.J. (1989). Patterns of pesticide use in New Zealand. Part 1. North Island, 1985-1988. DSIR, Hamilton, New Zealand.

Wilcock, R.J.; Close, M.E. (1990). Patterns of pesticide use in New Zealand. Part 2. South Island 1986-1989. DSIR, Hamilton, New Zealand.

Williamson, R.B.; Hickey, C.W.; Robertson, B. (2017). Preliminary Assessment of Limits and Guidelines Available for Classifying Coastal Waters. No. Auckland Council Technical Report 2017/035. Prepared by Diffuse Sources, NIWA and Wriggle Coastal Management for Auckland Council, Auckland. pp. 171.

Xu, E.G.; Chan, S.N.; Choi, K.W.; Lee, J.H.W.; Leung, K.M.Y. (2018). Tracking major endocrine disruptors in coastal waters using an integrative approach coupling field-based study and hydrodynamic modeling. *Environmental Pollution 233*: 387-394.

Appendix B Overview of methods used for monitoring state variables

State variables/ Methods	Sampling design	Sampling procedures	Laboratory analyses	Computational approaches and metrics	Thresholds, guideline values	Comments	References
DO	Sampling varies across the country (i.e., site extent, number of replicates, frequency and sampling time).	Measured in situ, automatic profilers, surface water	not-applicable	not-applicable	National guideline value available	Easy to measure with instrumentation, but requires ongoing maintenance. High frequency required as it can vary considerably over hours, days, seasons. Measurements are affected by salinity and temperature	Dudley et al. 2017
Sediment nutrients	See EMP, ETI Tool 2	See EMP, ETI Tool 2	At least two different analytes for nitrogen: see EMP (TKN, TP), ETI Tool 2 (TN, TP)	not-applicable	ETI Tool 2 - TN thresholds for eutrophication status	Spatially variable within estuary. Temporally variable with changes in freshwater inflows, nutrient cycling processes, resuspension, etc. Generally considered that historic levels may take a while to change. Links to ecosystem health status have been demonstrated for N There are also other measures that may be worth exploring e.g., pore water ammonium.	EMP - Robertson et al. 2002, ETI Tool 2 - Robertson et al. 2016

State variables/ Methods	Sampling design	Sampling procedures	Laboratory analyses	Computational approaches and metrics	Thresholds, guideline values	Comments	References
Sediment TOC	See EMP, ETI Tool 2	See EMP, ETI Tool 2	See ETI Tool 2	AFDW (ash free dry weight) - a surrogate measure can be converted to TOC, but conversions give highly variable results	ETI Tool 2 - thresholds for eutrophication status	Can be cheaply and directly measured. Spatially variable within estuaries Likely to show non-linear response to nutrient loading. Links to ecosystem health status have been demonstrated.	EMP - Robertson et al. 2002, ETI Tool 2 - Robertson et al. 2016
Sediment sulphides	ETI Tool 2 and various designs for subtidal surveys	Sediment TS and SCr in upper 2 cm of sediment (ETI Tool 2). The top 30 mm of one sediment core was analysed for total free sulphides (µM) (example from subtidal aquaculture consent monitoring).	Usually measured using calibrated probe, which can be difficult to use in the field. Must be analysed within hours of arrival at lab	Sediment TS and SCr in upper 2 cm of sediment.	ETI Tool 2 - thresholds for eutrophication status under development	Links to ecosystem health status have not been demonstrated, but likely to be directly linked at high concentrations.	EMP - Robertson et al. 2002, ETI Tool 2 - Robertson et al. 2016

State variables/ Methods	Sampling design	Sampling procedures	Laboratory analyses	Computational approaches and metrics	Thresholds, guideline values	Comments	References
Depth of RPD	Sampling varies across the country (i.e., site extent, number of replicates, frequency and sampling time).	Often measured on a sediment core using a ruler. Can be easily measured using an ORP probe and meter in situ. Oct-March.	not-applicable	not-applicable	ETI Tool 2 - thresholds for eutrophication status	Can be cost effective if visual method used, although this method does not always correspond with laboratory measures for sediments with high Fe (e.g., many west coast estuaries). Can be difficult to separate out effects of nutrients vs sedimentation event Spatially variable. Links to ecosystem health status demonstrated.	EMP - Robertson et al. 2002, ETI Tool 2 - Robertson et al. 2016
Broad scale extent of habitats	EMP, ETI Tool 2	EMP, ETI Tool 2	not-applicable	Actual area, % of intertidal, comparison with historical value	ETI Tool 2 - thresholds for eutrophication status	Can be easily measured Support from literature for relationships with values. Some links to ecosystem biodiversity status have been demonstrated.	EMP - Robertson et al. 2002, ETI Tool 2 - Robertson et al. 2016
Broad scale extent of dominant substrates	EMP, ETI Tool 2	EMP, ETI Tool 2	not-applicable	Actual area, % of intertidal, comparison with historical value	ETI Tool 2 - thresholds for eutrophication status	Can be easily measured Support from literature for relationships with values. Some links to ecosystem biodiversity status have been demonstrated.	EMP - Robertson et al. 2002, ETI Tool 2 - Robertson et al. 2016
Frequency of bathing beach closures	Linked to water Faecal Indicator Bacteria methods/design	See water FIB	not-applicable	TBD - this does not appear to be a widely used current metric,	See water FIB factsheet		Website 'https://www.lawa.org.nz/explore-data/swimming/'

State variables/ Methods	Sampling design	Sampling procedures	Laboratory analyses	Computational approaches and metrics	Thresholds, guideline values	Comments	References
				but could be determined using water FIB data or data from LAWA on swimmability			Report 'our-marine-environment.pdf'
Frequency of harvest closures	Linked to shellfish Faecal Indicator Bacteria methods/design	See shellfish FIB	not-applicable	Closures may be based on data other than water or shellfish FIB, such as river flows, rainfall, salinity, etc. No clear standardised approach - can be site specific	See water and shellfish FIB factsheets		TBD from MPI
Shellfish distribution and abundance	Replicated spatial surveys using cores or quadrats	See macrofauna	not-applicable	Examples given in MPI shellfish survey reports	None known		Example reports include 'MS1.7_5982514-FAR-2013-39-Distribution-and-abundance-of-toheroa' and 'MS1.7_5135190-FAR-2012-45-pipis-and-cockles-int-he-Northland-Auckland-and-Bay-of-Plenty-regions-2012'

State variables/ Methods	Sampling design	Sampling procedures	Laboratory analyses	Computational approaches and metrics	Thresholds, guideline values	Comments	References
Harvest area accessibility	TBC	TBC	not-applicable	None currently developed	None currently developed		NA
Finfish diversity and abundance	Study specific not standardised for SOE monitoring	Study specific not standardised for SOE monitoring	not-applicable	Study specific not standardised for SOE monitoring	None known		Reports 'MS1.7_Francis estuarine fishes nationwide.pdf' and 'MS1.7_Francis estuarine fishes nationwide suppl.pdf'

Appendix C List of online survey respondents

Project partners:

Anna Berthelsen (Cawthron)

Anastasija Zaiko (Cawthron)

Candida Savage (UO)

Chris Cornelisen (Cawthron)

Chris Hickey (NIWA)

Conrad Pilditch (UoW)

Darren Parsons (NIWA)

Graham McBride (NIWA)

Jonathan Banks (Cawthron)

Judi Hewitt (NIWA)

Leigh Stevens (Wriggle/ Salt Ecology)

Megan Carbines (AC)

External technical experts:

Andrew Swales (NIWA)

Chris Hickey (NIWA)

Jennifer Gadd (NIWA)

John Zeldis (NIWA)

Louis Tremblay (Cawthron)

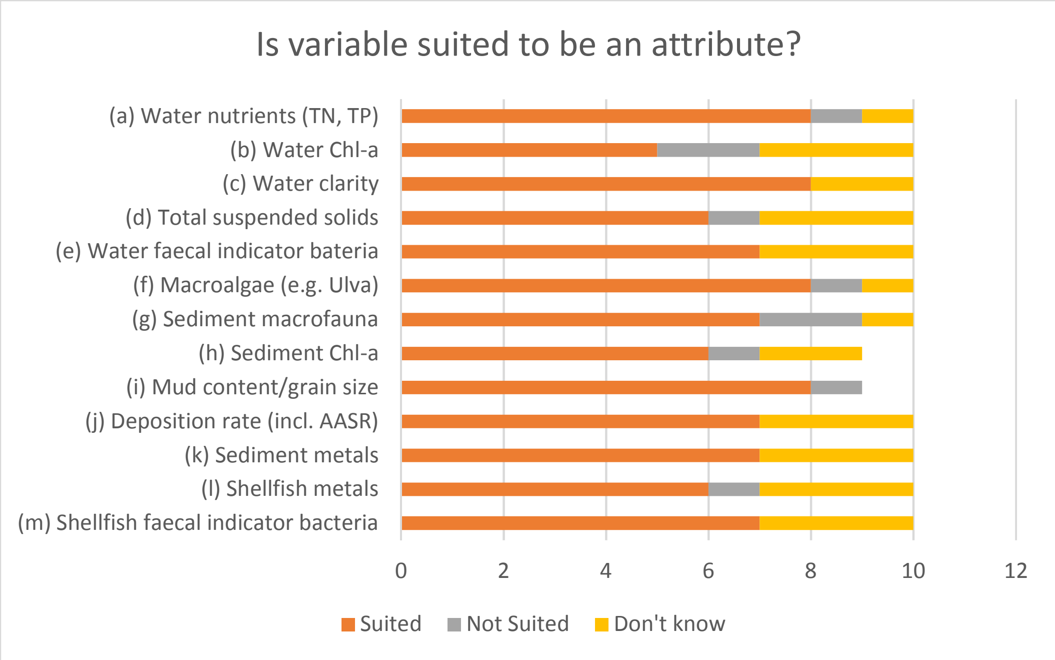
Max Gibbs (NIWA)

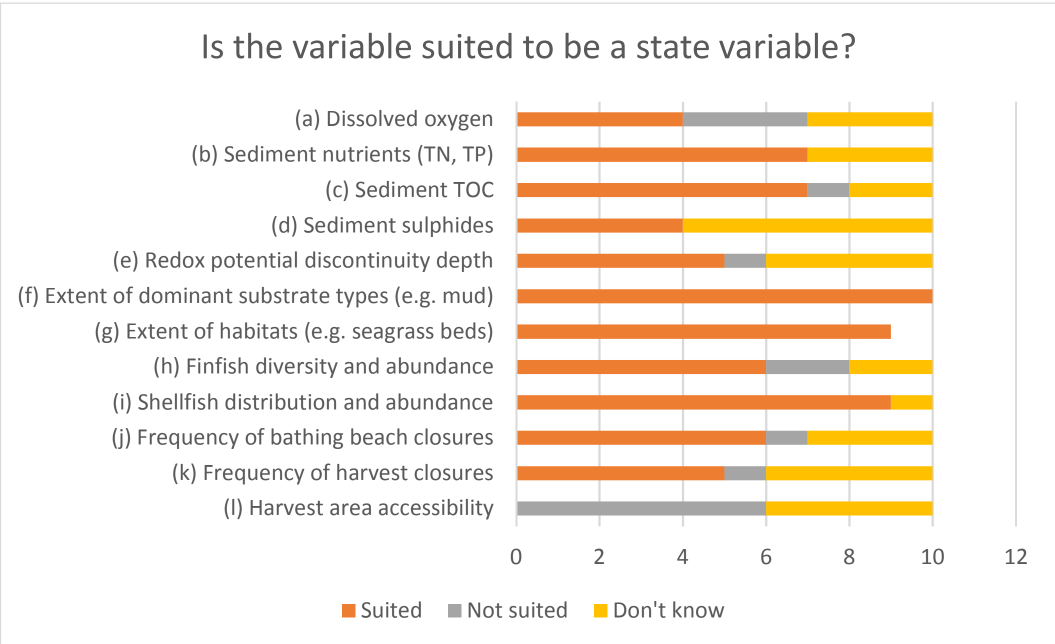
Paul Gillespie (Cawthron)

Rob Davies-Colley (NIWA)

Stephen Park (BOPRC)

Appendix D Results of the online survey Part A





Appendix E Lag and recovery consideration for prioritised attributes and state variables.

Table E-1: Indicative lag and recovery timscales for prioritised attributes and state variables. Lag categories: D – Days; W – Weeks; M – Months; Y – Years.

Attribute/ State variable	Signal detection	Recovery	Comments
Water nutrients (TN, TP)	D	D-W-Y	An initial change would be detected fairly rapidly, but dependent on estuary residence time and how much comes back on the next time it could takes days to finally flush out. It is also important to remember that much of the nutrients found in the water column in an estuary can be provided by fluxes from the sediment. Nutrient uptake by macroalgae and saltmarsh may affect measured concentrations
Water Chl-a	D-W	D-W-Y	Will be delayed relative to TN-TP in water column. Retention is the primary driver as will be flushed from estuary if not trapped by stratified waters or mouth closure or constriction
Water clarity	D-M	W-Y	Due to potential for resuspension of sediments water clarity changes would probably only be detected in water entering the estuary quickly, the rest of the estuary would take weeks to years to change
Total suspended solids	D-M	W-Y	As above
Water faecal indicator bacteria	D	D-W-M?	Dependent on estuary residence time and influence of solar disinfection
Macroalgae	W-M	M-Y	Cannot change as fast as macrofauna can, as not mobile. Still, can grow rapidly in response to available nutrients and under suitable growing conditions. Can also be uprooted and flushed from estuary
Macrofauna	D-W	W-Y	Most macrofauna would be able to start changing rapidly as soon as TSS and sedimentation rate decrease although full recovery would be dependent on removal of the built up sediment or contaminants
Sediment Chl-a	D-W	W-Y	As above
Mud content/grain size	Y	Y	Would require resuspension and flushing out of the estuary to remove
Deposition rate (incl. AASR)	D	W-Y	Due to potential for resuspension of sediments AASR changes would probably only be detected in sedimentary environments of the upper estuary, the rest of the estuary would take weeks to years to change

Attribute/ State variable	Signal detection	Recovery	Comments
Sediment metals and emerging contaminants	M-Y	Y	Less time than mud content as new uncontaminated sediment can dilute present sediment. On the other hand, fresh contaminated sediment may overly clean sediment and be quicker to detect
Shellfish metals and emerging contaminants	D	D-Y	This is for metals- signal could take longer for some contaminants
Shellfish faecal indicator bacteria (or a virus or virus indicator)	D-W	D-W-(M?)	Signal could take longer for some indicators and pathogens. Possible to have FIB absent but pathogens still remain in shellfish
Dissolved oxygen	M-Y	Y	Assuming this is in the sediment, for water see comments for water TN/TP
Sediment nutrients (TN, TP)	M-Y	Y	As per sediment contaminants
Sediment TOC	M-Y	Y	As per sediment contaminants
Sediment sulphides	M-Y	Y	As per sediment contaminants
Redox potential discontinuity depth	M-Y	Y	
Extent of dominant substrate types	Y	Y	
Extent of habitats	M-Y	Y	
Finfish diversity and abundance	D-M	W-Y	Dependent on timing of change, recruitment or species movement into estuary
Shellfish distribution and abundance	D-W	W-Y	for juveniles- for adults both would be Y
Frequency of bathing beach closures	D	D	Depends on follow-up sampling
Frequency of harvest closures	D-W	D-W	Depends on follow-up sampling
Harvest area accessibility			