National Estuary Dataset: User Manual

OTOT Report No.5
April 2018

Oranga Taiao Oranga Tangata

Report No. 5

2018

National Estuary Dataset:

User Manual

National Estuary Dataset:

User Manual

Anna Berthelsen[1]
Dana Clark[1]
Eric Goodwin[1]
Javier Atalah[1]
Murray Patterson[2]

[1] Cawthron Institute, Nelson New Zealand
[2] School of People, Environment and Planning, Massey University, Palmerston North, New Zealand

Published by the Oranga Taiao Oranga Tangata (OTOT) Research Team
Funded by the Ministry for Business, Innovation and Employment
Contract MAUX1502
Contract Holder: Massey University
www.mtm.ac.nz

Reviewed by: Jim Sinner

Approved for release by: Manager, Coastal Group
Chris Cornelisen

Recommended citation
Berthelsen A., Clark D., Goodwin E., Atalah J., Patterson M. (2018). National Estuary Dataset: User Manual. OTOT Research Report No. 5. Cawthron Report No. 3152. Massey University, Palmerston North.

Published by the Oranga Taiao Oranga Tangata (OTOT) Research Team
Cawthron Report No. 3152
Contract Number MAUX1502
Contract Holder: Massey University
Private Bag 11052
Palmerston North
New Zealand

ISBN 978-0-9951033-9-9 (Print)
ISBN 978-0-9951033-8-2 (Online)

Table of Contents

List of Acronyms

AC = Auckland Council

ADL = analytical detection limit

AFDW = ash-free dry weight

As = arsenic

BOPRC = Bay of Plenty Regional Council

CCC = Christchurch City Council

Cd = cadmium

cesym = council estuary site year month

CMEC = Coastal Marine Ecology Consultants

Cr = chromium

CRC = Canterbury Regional Council

Cu = copper

DOI = digital object identifier

ECAN = Environment Canterbury

ECHI = Estuarine Cultural Health Index

EMP = Estuary Monitoring Protocol

EOS = EOS Ecology

ES = Environment Southland

GIS = geographic information systems

GWRC = Greater Wellington Regional Council

HBRC = Hawkes Bay Regional Council

ISPT = Integrative Spatial Planning Tool

LOI = loss on ignition

MBIE = Ministry of Business, Innovation and Employment

MDC = Marlborough District Council

MTM = Manaaki Taha Moana

NA = not available

NCC = Nelson City Council

Ni = nickel

NIWA = National Institute of Water and Atmospheric Science

NRC = Northland Regional Council

NZTME = New Zealand Transverse Mercator Easting

NZTMN = New Zealand Transverse Mercator Northing

ORC = Otago Regional Council

OTOT = Oranga Taiao Oranga Tangata

Pb = lead

QA = quality assurance

RPD = Redox Potential Discontinuity

Ryder = Ryder Consulting

TDC = Tasman District Council

TKN = total Kjeldahl nitrogen

TOC = total organic carbon

TN = total nitrogen

TP = total phosphorus

Triplefin = Triplefin Environmental Consulting

WRC = Waikato Regional Council

WCRC = West Coast Regional Council

WoRMS = World Register of Marine Species

Zn = zinc

1. Introduction

Cawthron Institute (Cawthron) has recently compiled a national dataset containing ecological estuary monitoring data (2001 to 2016) largely acquired from regional councils and unitary authorities[1] around New Zealand. The dataset comprises fine-scale intertidal benthic ecological data collected using the Estuary Monitoring Protocol (EMP; Robertson et al., 2002), or similar survey methodologies. This is in the form of macrofaunal abundance data and corresponding physico/chemical sediment data, as well as associated metadata.

The dataset was compiled to facilitate national-scale research within the MBIE-funded *Oranga Taiao, Oranga Tangata* (OTOT) programme[2] (refer Section 2 for more details). Within the OTOT programme, we have already used a subset of the dataset to test the performance of biotic indices of estuary health (Berthelsen et al., 2018). Future use of the dataset is planned within the OTOT programme, and it will likely be useful for others as well. This report aims to assist users by providing a 'user manual' to accompany the dataset. It includes details of the dataset relating to the following:

- overview of data
- standardised coding for each unique sampling event
- sampling design
- sample collection methodology
- laboratory analytical methodology
- quality assurance
- data management.

A report detailing inconsistencies in the data and the issues these caused for compilation and analysis has recently been published (Berthelsen, Atalah, & Clark, 2017). As that report and the current report were generally written to be independent of one another (i.e. stand-alone), some information is included in both reports. However, users of the dataset may find both to be of interest.

[1] A territorial authority (district or city) which also performs the functions of a regional council.
[2] https://www.mtm.ac.nz/oranga-taiao-oranga-tangata/

2. Oranga Taiao Oranga Tangata

The National Estuary Dataset was compiled for the MBIE-funded programme *Oranga Taiao, Oranga Tangata: Knowledge and Toolsets to Support Co-Management of Estuaries* (MAUX1502) which builds upon a previous MBIE-funded programme, Enhancing Coastal Ecosystems for Iwi: Manaaki Taha Moana (MAUX0907). The OTOT research programme ($4.4 million + GST) has a case study that focuses on the Tauranga Harbour and its catchment. It is a four-year research programme (October 2015 to September 2019) that has three phases.

Phase 1 focuses on gathering Mātauranga Māori (a body of knowledge of Māori experience in the area) from local iwi/hapū. From this information, an Estuarine Cultural Health Index (ECHI), or other similar tool(s), will be constructed so that iwi/hapū can assess the state of local estuarine habitats, record changes over time and help judge the effectiveness of factors such as local fishing rules and management strategies.

Phase 2 will consolidate the ecological knowledge of the Tauranga Harbour and begin to provide some modelling and indicators of estuarine ecosystem health, resilience and functioning.

Phase 3 will see the creation of an Integrative Spatial Planning Tool (ISPT). This tool is a hybrid Graphic Information System (GIS)/modelling system that will use information from the estuarine ecology, land use, economic and cultural areas, where appropriate. It will enable users to evaluate future planning options for Tauranga Harbour. This integrative (ecological, economic, land use, cultural, demographic) planning tool should be at the leading edge of developments worldwide. Although such tools have been developed for the terrestrial environment, few if any spatial-modelling tools have been developed for the-whole-of catchment including both land and coastal-marine ecosystems.

In all phases, the knowledge, frameworks and toolsets developed will be developed in such a way to foster transference and uptake to other iwi and regions throughout New Zealand, where possible, to enhance the health of estuaries nation-wide, and indeed internationally.

3. Overview of dataset

We derived the National Estuary Dataset (Clark et al., 2018) from fine-scale intertidal benthic ecological data collected using the EMP (Robertson et al., 2002), but also included data from similar survey methodologies. Although most of the data were collected by councils for the purpose of State of the Environment monitoring, the dataset also includes some consent monitoring data from Porirua Harbour (Boffa Miskell, 2014)[3] in the Wellington region, research data collected for the Manaaki Taha Moana programme from Tauranga Harbour in the Bay of Plenty region (Ellis et al., 2013)[4], and data collected for the development of the EMP (Robertson et al., 2002)[5] from seven regions nationally (Northland, Bay of Plenty, Tasman, Marlborough, Canterbury, Otago and Southland). Although these additional data were not collected by councils, in the dataset (and throughout this report) we have, for simplicity, used council names to define regions from which data were acquired. For example, the research data from the Tauranga Harbour survey are labelled as Bay of Plenty Regional Council (BOPRC) even though they were collected by researchers (although the council assisted with the survey).

The raw data were acquired from the regions of fourteen councils and the dataset contains information from 70 estuaries, 409 sites and 815 sampling events (Table 1, Figure 1). Data were not able to be acquired from some councils, e.g. Gisborne District Council, Taranaki Regional Council, Horizons Regional Council, or from other sources for their regions.

The dataset contains intertidal (but no subtidal) macrofaunal abundance data (sieved through 0.5 mm mesh, with all sieved taxa included) and corresponding sediment physico/chemical data for at least one (but ideally all) of the following variables:

- grain size
- nutrients
- organic content
- metals
- associated metadata.

The data were usually acquired as a Microsoft Excel spreadsheet (raw data file).

We largely relied on obtaining metadata from the raw data files and reports, and only emailed key council contacts if we could not find the information in the files and reports.

Although we aimed to acquire and then include all available data that met our requirements, the dataset does not necessarily contain all data collected for ecological estuarine monitoring programmes during this period. Some data that met the criteria above were deliberately not included. For example, Auckland Council (AC) data prior to 2010 were not included in the dataset

[3] All data from the estuary Porirua from the years 2013 and 2014.
[4] All data in the dataset from the estuary Tauranga.
[5] All data in the dataset from the year 2001.

as it was recognised that macrofaunal taxonomic identification was conducted at a lower resolution (Ebrahim Hussain, Auckland Council, pers. comm.). Some data met our criteria but have unintentionally not been included in the dataset at this stage. The example we know of is some of the more recent data from Northland Regional Council (NRC) sentinel sites. We also chose to exclude all data for some variables e.g. macroalgal cover, epifauna abundance and sediment chlorophyll-a, phaeophytin, organic compounds and Redox Potential Discontinuity (RPD) depth, due to inconsistencies in sampling frequency, methodology sample collection and analysis and/or data availability.

Table 1. Number of estuaries, sites, sampling events and years included for each council in the National Estuary Dataset.

Council	No. of estuaries	No. of sampling events	No. of sites	Years	First year	Last year
Auckland Council (AC)	13	219	93	5	2010	2014
Bay of Plenty Regional Council (BOPRC)[6]	2	78	78	2	2001	2011
Environment Canterbury (ECAN)	4	34	8	8	2001	2015
Environment Canterbury (ECAN)/Christchurch City Council (CCC)[7]	1	43	7	7	2007	2015
Environment Southland (ES)	8	65	23	12	2001	2013
Greater Wellington Regional Council (GWRC)	9	74	34	9	2004	2014
Hawke's Bay Regional Council (HBRC)	4	54	8	10	2006	2015
Marlborough District Council (MDC)	5	16	12	4	2001	2016
Nelson City Council (NCC)	2	6	6	2	2009	2012
Northland Regional Council (NRC)	8	105	99	8	2001	2016
Otago Regional Council (ORC)	8	17	17	6	2001	2012
Tasman District Council (TDC)	3	24	9	5	2001	2015
Waikato Regional Council (WRC)	3	78	13	2	2013	2014
West Coast Regional Council (WCRC)	1	2	2	1	2007	2007

The raw data varied widely in reporting format, reporting conventions for variable names, site identifiers, date formats, units of measurement, and other data structure elements. We imported

[6] Research data only – not from the council's estuary monitoring programme.
[7] Data and metadata for the same sampling events were provided by both councils.

the datasets into the statistical software program 'R' and imposed a consistent set of reporting conventions. Aligning macrofaunal data, sediment physico/chemical data and associated metadata was an intensively controlled process, coordinated by the "cesym" identification code described in Section 4 below. Each row in the dataset represented a single sampling event (i.e. a sampling occasion where variables were measured concurrently at the same site).

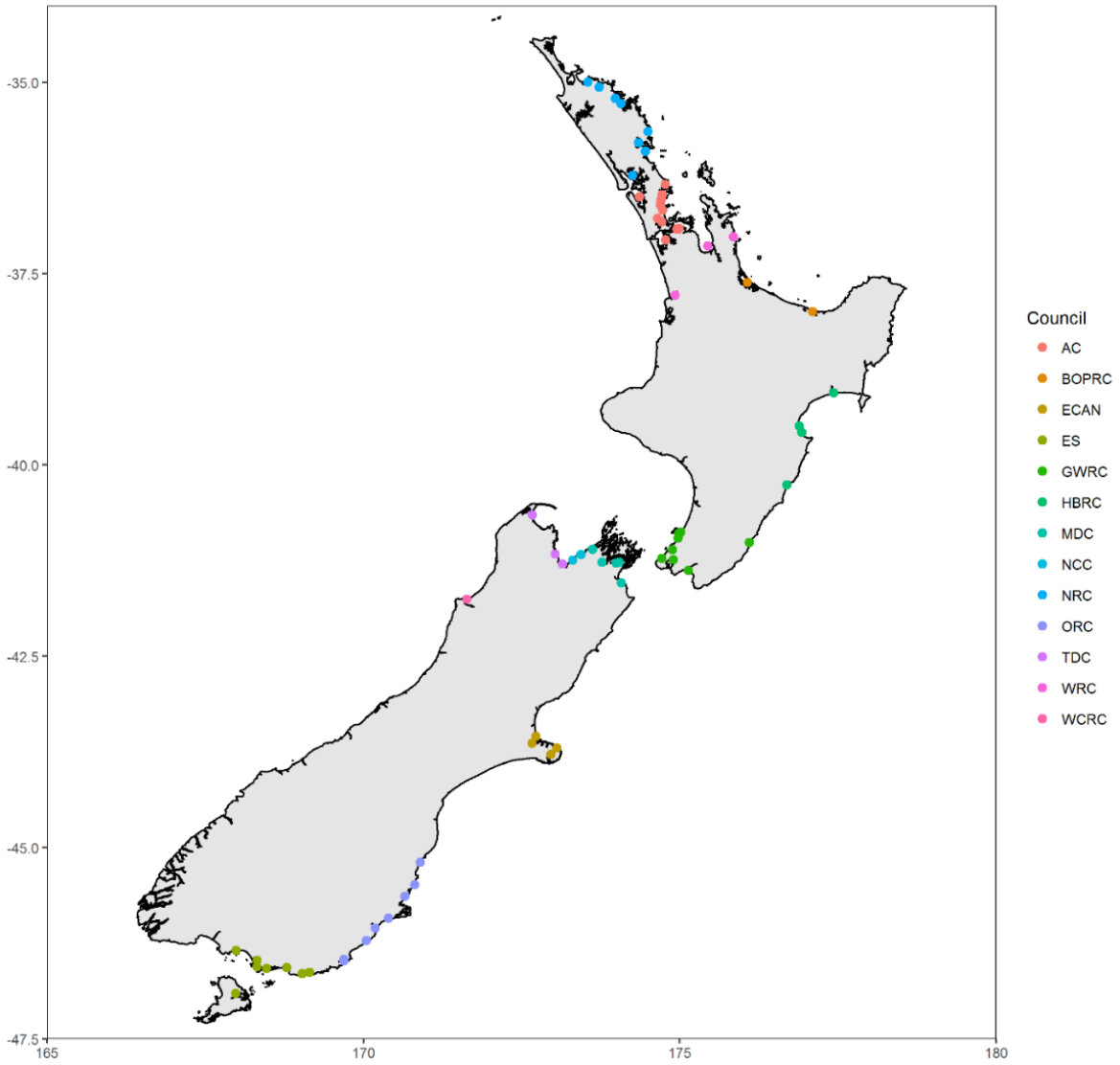

Figure 1. Map showing the geographic locations (dots colour-coded to council) of all estuaries in the National Estuary Dataset.

4. Standardised Coding

Standardised coding was used to identify individual sampling events within the dataset. Each sampling event can be identified by a unique code specifying its council_estuary_site_year_month referred to as cesym (code terms described in Table 2) e.g. aucklandregionalcouncil centralwaitemata_hbv_2010_october. If only one sampling event was undertaken within a given year, the specific month was replaced with the word 'all' (e.g. westcoastregionalcouncil_orowaiti a_2007_all). A separate column in the dataset specifies the actual month during which each sampling event was conducted. As there is some duplication in estuary and site names, the full cesym code is needed to identify each individual sampling event. For example, there is a Waitangi Estuary in both Northland and Hawkes Bay, and many councils use a simple numbering or lettering system to assign site names. If the user would like to count the number of estuaries and sites, or average the data at the level of site or estuary, this needs to be accounted for. Replicates within a sampling event (i.e. cesym) can be further identified using the code council_estuary_site_year_month_replicate (cesymr).

Table 2. Terms used for standardised coding to identify individual sampling events within the National Estuary Dataset. The information was sourced from raw data files, relevant reports and communication with key council contacts.

Column name in dataset	Description	Comment	Example from dataset
council	Council name	The council region within which the data were collected.	aucklandregionalcouncil
estuary	Estuary name	Name of the estuary	centralwaitemata
site	Site name	Name of the site	hbv
year	Year of sampling event	Year during which the sampling event was conducted	2010
month	Month of sampling event	The actual name of the month was only used in the cesym code if more than one sampling event was conducted during a year, otherwise denoted as 'all"	October
replicate	Replicate name	Name of each replicate for macrofaunal data within a sampling event, usually a number or letter	1
cesym		Standardised code to identify individual sampling events	aucklandregionalcouncil centralwaitemata hbv 2010 october
cesymr		Standardised code to identify unique replicates within an individual sampling event	aucklandregionalcouncil centralwaitemata hbv 2010 october 1

5. General Sampling Design

Sampling events were largely conducted following the sampling design described in the EMP (Robertson et al., 2002), although there was some variation (e.g. site size and location in terms of representativeness, replicate number - including compositing for physico/chemical samples).

The maximum area (i.e. site size), within which all macrofaunal and physico/chemical sediment samples were collected during a sampling event was 10,800 m^2 (Halliday, Townsend, & Lundquist, 2012) although in most cases this was considerably smaller e.g. EMP specifies a site size of 1800 m^2 (Robertson et al., 2002). Due to the time required to obtain this information (e.g. by searching through relevant reports and/or communicating with councils), the specific site size for each sampling event was not included as metadata within the dataset.

Overall, we considered the different sampling designs to be comparable. However, we included metadata describing the tidal height, vegetation cover and location of sites (Table 3) and number of macrofaunal replicates (Table 2), so these factors could be considered as part of the analyses if necessary. No further information is provided in the dataset in terms of variation between sampling designs, although this can be further investigated by future users of the dataset by querying relevant reports (where these exist) or by communicating directly with councils.

Table 3. Metadata (site description, location and sampling month) associated with sampling design within the National Estuary Dataset. Note that NA indicates information that was not available at the time. The information was sourced from raw data files, relevant reports and communication with key council contacts.

Column name in dataset	Categories	Description
tidal.height	low mid/low mid mid/high NA	General height of sampling site in relation to the tide. Note that the sites in the mid/low category could belong to either the low or mid categories, however at the time of data compilation this was unknown.
vegetated.unvegetated	vegetated unvegetated NA	Description of whether a site was vegetated (i.e. covered with seagrass, mangroves or macroalgae) or unvegetated. Note that in some cases unvegetated sites contained small amounts of macroalgae.
vegetated.detail.unvegetated	mangrove seagrass seagrass/macroalgae seagrass/mangrove unvegetated NA	If site was considered vegetated (see row above), further description was given as to what type of vegetation. If it was unvegetated, it was also given the category unvegetated in this column.
month.x	January, February etc. NA	Month during which sampling was conducted
NZTME		General location of sampling site in New Zealand Transverse Mercator 2000 (NZTM2000) coordinates. In a small number of cases where coordinates were not available this was estimated from images of site locations.
NZTMN		General location of sampling site in New Zealand Transverse Mercator 2000 (NZTM2000) coordinates. In a small number of cases where coordinates were not available this was estimated from images of site locations.

6. Macrofauna

Sample collection and analysis

Macrofaunal samples were collected by pushing cylindrical cores into the sediment and sieving the contents through a 0.5 mm mesh sieve. In most cases cores were 130 mm in diameter and pushed into the sediment to a depth of 150 mm, but in a few cases cores with diameters of 125 or 150 mm were used, or cores only pushed to 100 mm depth (Table 4). All macrofaunal individuals were identified to the lowest taxonomic level practicable by a variety of taxonomic experts throughout the country.

Macrofaunal data were kept at the replicate (i.e. core) level within the dataset (as described in Section 5, Table 2) and was represented by the abundance of each individual taxa per replicate. We scaled down and up abundances of each taxa in the 150 and 125 mm diameter cores respectively, based on the proportional difference of each diameter from 130, to standardise with the 130 mm diameter cores. Therefore, macrofauna abundances are reported in terms of counts per core surface area rather than counts per volume[8].

[8] Raw counts were standardised by dividing raw counts by the diameter of the macrofaunal core used the collect the samples and multiplying by a standard 130 mm. This means the values for all macrofaunal taxa are in units of counts per 130 mm core.

Table 4. Metadata associated with sample collection and analysis of macrofaunal data within the National Estuary Dataset. The information was sourced from raw data files, relevant reports and communication with key council contacts.

Column name in dataset	Categories	Units (if applicable)	Description
core.diameter..mm.	125 130 150	mm	Diameter of core used to collect macrofaunal samples for a given sampling event
core.depth..mm.	100 150	mm	Depth of core used to collect macrofaunal samples for a given sampling event
taxonomy.by.	Boffa Miskell[a] Cawthron[b] EOS[c] CMEC[d] CRC[e] NIWA[f] NRC/CMEC[g] Ryder[h] Triplefin[i] WRC[j] Wilma Blom[k]	-	Name of organisation/company or taxonomist who conducted taxonomic analysis for a given sampling event

[a] *Boffa Miskell http://www.boffamiskell.co.nz/*
[b] *Cawthron Institute http://www.cawthron.org.nz/*
[c] *EOS Ecology http://www.eosecology.co.nz/*
[d] *Coastal Marine Ecology Consultants (principle Gary Stephenson)*
[e] *Canterbury Regional Council - Lesley Bolton-Ritchie, coastal water quality and ecology scientist.*
[f] *National Institute of Water and Atmosphere https://www.niwa.co.nz/*
[g] *Northland Regional Council and CMEC (principle Gary Stephenson) – macrofauna were largely sorted and identified by NRC staff but small and/or cryptic fauna were sent to CMEC.*
[h] *Ryder Consulting http://www.ryderconsulting.co.nz/*
[i] *Triplefin Environmental Consulting https://www.triplefin.co.nz/*
[j] *Waikato Regional Council – Nathan Singleton*
[k] *Wilma Blom - curator of marine invertebrates at Auckland War Memorial Museum*

Merging macrofaunal data

There were a variety of issues of inconsistency in taxonomic naming between the raw data files including the presence of synonyms, misspellings, species codes (e.g. Polychaete sp. A) and common names (e.g. tuatua). We followed the World Register of Marine Species (WoRMS Editorial Board, 2017) for taxonomic nomenclature. Considerable effort was made in making taxonomic descriptors consistent in data files obtained from different councils or other sources, in R (R Core Team, 2017) using the library taxize (Chamberlain & Szocs, 2013) and (Chamberlain et al., 2016), and taxizesoap (Chamberlain) packages to query the online WoRMS database.

We retained juveniles in separate columns in the dataset using the code 'taxon name juvenile' (e.g. maldanidae juvenile) where they were identified separately by taxonomists. This allows future users to make their own decisions regarding how to treat these in the data e.g. remove, or keep separate from or combine with parent taxa. We note that whether or not juveniles were recorded separately from their parent taxa appeared to be inconsistent across the raw data files. In some cases, size classes had been recorded for certain bivalve taxa and we lumped all size classes together in the dataset without trying to differentiate juveniles.

We did not use the terms sp. or spp. in the dataset, so any taxon identified at a level higher than species can include one or more taxa. All vertebrates (e.g. fish), plants (e.g. macroalgae), bacteria, and larval planktonic groups (e.g. megalope, larvae, eggs) were removed. Taxa that traditionally may not be considered macrofauna (e.g. Porifera, Tunicata, Ascidiacea, Bryozoa, Daphnia and Insecta) were retained to allow users to decide whether to remove these or not prior to analysis. Higher level taxonomic information for each taxon has been included in a separate dataset to aid the implementation of these decisions (File name: Higher_Level_Taxonomic_Information _Final2017-11-10.csv). Zero abundance for a taxon in a replicate was indicated by a zero value in the dataset.

Taxonomic lumping

Shade plots of the presence/absence of taxa were created in the statistical programme PRIMER 7 (Clarke, Gorley, Somerfield, & Warwick, 2014) to detect differences in the level of taxonomic resolution between data analysed by different taxonomists. The plots indicated that lumping of taxonomic groups was required to increase data comparability across the dataset. To allow users to make their own decisions regarding taxonomic resolution, no lumping of taxa (besides that required for the initial cleaning/grooming of the data e.g. resolving synonyms) was conducted in the dataset. However, we strongly recommend that some lumping of taxa is undertaken before data analysis to ensure comparability across the dataset. We have suggested an option for lumping in Appendix A.

7. Physico/chemical sediment data

Sample collection

Physico/chemical sediment samples were generally collected using EMP methodology with some variation (e.g. samples collected within a grid versus randomly within a site). In a small number of cases, sampling of metals was not concurrent with sampling of other variables, and was instead collected on a slightly different date (e.g. all metals data from Waikato Regional Council - WRC). The number of physico/chemical replicates analysed per sampling event ranged from one to twelve. This variation in replicate number arose from differences in sampling effort and/or compositing of samples prior to laboratory analyses in some surveys, resulting in a lower number of replicate samples than originally collected.

Laboratory analyses

Sediment samples were analysed for grain size, nutrients and metals (Table 5), although not all variables were measured during each sampling event. Laboratories that conducted the analyses included: Auckland UniServices, Cawthron Institute, Hill Laboratories, National Institute of Water and Atmosphere (NIWA), University of Waikato and Watercare Laboratory Services. Although the EMP recommends that sampling sites have overlying water with salinity > 20 ppt (Robertson et al., 2002), this information was generally not available and therefore not included in the dataset.

Total phosphorus and the two main measures of organic content (ash-free dry weight–AFDW and total organic carbon–TOC) are displayed in separate columns in the dataset with no associated metadata. Nitrogen and categories for each grain size (with recalculation as required) are also displayed in one column each, however metadata is provided to discriminate between various laboratory analysis methods (Table 6) to allow future users to make their own decision regarding the comparability of different methods.

Sediment grain size was analysed using two main methods: laser diffraction analysis and wet sieving, and results from these are not necessarily comparable (Bolton-Ritchie & Lawton, in draft; Hewitt, Hailes, & Greenfield, 2014; Mills & Williamson, 2014). Some variation may also exist within these methodologies (Bolton-Ritchie & Lawton, in draft; Appendix B), although metadata to identify this variation was not included in the dataset. As grain sizes were often reported in different size classes, we recalculated these to form three size classes (< 63 µm, 63 µm–2 mm, > 2 mm). To increase comparability between different sediment grain size analyses, we converted sediment proportions per size class to a percentage of the 2-mm sediment fraction (e.g. percentage of < 63 µm out of the < 2 mm sediment fraction), although we also kept the original values (Table 5). This is because the maximum grain size analysed differed between analysis methods e.g. Malvern Mastersizer (laser) only analyses grains < 2 mm, while all grain sizes are generally analysed during wet sieving. Nitrogen was analysed as either total nitrogen (TN) or Total Kjeldahl Nitrogen (TKN).

Metals data are displayed in one column per metal under the assumption that analysis from the < 500 µm or < 2 mm/total sediment grain size fractions and variation in methods (Appendix B) gave comparable results.

Data merging

To merge the raw physico/chemical data files into one overall dataset we first used the R software programme (R Core Team, 2017) to group together variables assumed to be the same (e.g. AFDW and loss on ignition–LOI). We plotted density distributions of each variable in each group, and visual comparison of the plots indicated when data from different sources had been reported in different units (e.g. mg/kg *vs* g/100 g, etc) where this was otherwise unclear (i.e. unit details not provided). Conversions were made if required.

We then averaged the replicate values for each physico/chemical variable per sampling event (cesym) i.e. the same variable value (average) was assigned to each replicate (based on macrofaunal data) within a sampling event. This was because paired replicates for macrofauna and physico/chemical variables were not always collected and, even if they were, compositing of samples in some cases obscured the relationships between paired samples. The number of replicates for each sampling event were not provided as metadata.

All variables (based on average variable value per sampling event) below the Analytical Detection Limits (ADL) were replaced with zero values. ADLs for some variables differed across laboratories and it was often unknown whether these had been previously adjusted in the raw data files e.g. use of the common convention of substituting ADL values with half of the ADL. Therefore, our rule for all variables, except nitrogen, was to apply the highest ADL known for each variable to all sampling events (Table 5). The nitrogen ADL for some laboratories was particularly high in relation to possible ecological impacts. For example, the ADL for TN analysed by Hill Laboratories was 500 mg/kg, however TN at concentrations of 250-1000 mg/kg causes minor stress to sensitive organisms (interim threshold only; Robertson et al., 2016). To avoid unnecessarily replacing TN values with zero, we applied different ADLs for nitrogen depending on the laboratory that conducted the analysis.

Within the dataset NA associated with the physico/chemical data and metadata indicates that the data was either not available (e.g. either not collected during time of sampling or not provided to us), or not applicable (e.g. for nitrogen type where nitrogen was not measured during sampling).

Table 5. Sediment physico/chemical variables within the National Estuary Dataset. The information was sourced from raw data files, relevant reports and communication with key council contacts and laboratories.

Column name in dataset	Variable category	Unit	Description	ADL applied
sedlt63	Sediment grain size	%	% sediment < 63 µm of total analysed (total from which fraction is analysed can differ for different analysis types)	None
sed63umto2mm	Sediment grain size	%	% sediment 63 µm-2 mm of total analysed (total from which fraction is analysed can differ for different analysis types)	None
sedgt2mm	Sediment grain size	%	% sediment > 2 mm of total analysed (this only applies to wet sieving as laser does not analyse grains > 2 mm)	None
sedlt2mm	Sediment grain size	%	% sediment < 2 mm of total analysed	None
sedlt63OOlt2mm	Sediment grain size	%	% sediment < 63 µm of total sediment < 2 mm	None
sed63to2OOlt2mm	Sediment grain size	%	% sediment 63 µm-2 mm of total sediment < 2 mm	None
TOC	Organic content	g/100g	Total Organic Carbon	0.05 g/100g
AFDW	Organic content	g/100g	Ash Free Dry Weight	0.04 g/100g
Cu	Metal	mg/kg	Copper	2 mg/kg
Cr	Metal	mg/kg	Chromium	2 mg/kg
Zn	Metal	mg/kg	Zinc	7.5 mg/kg
Ni	Metal	mg/kg	Nickel	2 mg/kg
Pb	Metal	mg/kg	Lead	1 mg/kg
Cd	Metal	mg/kg	Cadmium	0.1 mg/kg
As	Metal	mg/kg	Arsenic	2 mg/kg
TN	Nutrient	mg/kg	Nitrogen	250 mg/kg for all values from sampling events conducted in 2001. 50 mg/kg for all values from sampling events conducted by Northland Regional Council, except for those from 2001.

Column name in dataset	Variable category	Unit	Description	ADL applied
				500 mg/kg for values from all other sampling events not described above.
TP	Nutrient	mg/kg	Total/Total Recoverable Phosphorus	40 mg/kg

Table 6. Metadata associated with sediment physico/chemical variables within the National Estuary Dataset. The information was sourced from raw data files, relevant reports and communication with key council contacts.

Metadata column name in dataset	Category overview	Description
grain.size.method	laser	laser = laser diffraction
	wet sieve	wet sieve = wet sieving
nitrogen.type	TKN	TKN = Total Kjeldahl Nitrogen
	TN	TN = Total Nitrogen
	NA	NA = nitrogen was not measured at all during the sampling event

8. Quality Assurance

We have not provided information regarding the quality of the information in the raw data files and associated reports. We also note that the difficulties in obtaining some metadata possibly increased the chance of this information being inaccurate (Berthelsen et al., 2017). However, we did conduct quality assurance (QA) procedures on the National Estuary Dataset to help ensure it accurately reflected the raw data. The QA procedure was implemented by comparing the raw data values against randomly selected cesyms (including all associated replicates) for the following data:

- abundance of three macrofaunal taxa
- values for all physico/chemical variables
- metadata information.

Initially twenty cesyms (including all associated replicates) were QA'd following the above procedure; any issues identified were resolved in an updated dataset. After this another eleven cesyms were QA'd and all issues resolved, and then after this another ten cesym's were QA'd for which we got a pass rate of 100 percent. Overall, five percent of cesym's were put through the QA process.

The QA procedure also included an accuracy check (and update if require) of the:

- highest and lowest values, as well as any obvious anomalies, for each physico/chemical variable.
- taxonomic list for missing taxa, inconsistent naming of taxa, and any taxa with zero abundance values.

The QA process was implemented to provide some certainty regarding accuracy of the dataset. However, use of the dataset is entirely at the risk of the recipient and Cawthron accepts no responsibility for any inaccuracies that may be present.

9. Dataset management

The National Estuary Dataset is deposited in Figshare, an online data repository (www.figshare.com). Figshare is a cloud-based data repository where researchers and institutions can upload and store data. A DOI (digital object identifier) is created for deposited databases, as a persistent citable link, which can be used to reference the data. The National Estuary Dataset has been deposited as a confidential file in a Cawthron Figshare account that will allow OTOT to maintain control of who can use it (Clark et al., 2018). Permission to use the National Estuary Dataset must be gained from the OTOT programme[9].It is intended that the dataset will become publicly available once the OTOT research programme is completed in 2020.

Use and copyright of the dataset will be governed by an Attribution 4.0 International Creative Commons licence (CC BY 4.0, www. creativecommons.org/licenses/by/4.0/). Under this licence users can copy, and share the dataset; as well as adapt, transform, and build upon the dataset for any purpose, even commercially, as long the source is attributed by citation, a link to the license is provided, and any changes made to the database are indicated.

It is envisioned that additional raw data will be added to the National Estuary Dataset in the future as it becomes available. The details within this report should be used as a guide for this process with emphasis on the following actions:

- ensure all data is comparable (consider sampling design/collection, sample analysis – may need to convert units and grain size, apply designated ADLs, metadata, ensure taxon names and resolution are the same, scale macrofauna abundances if core diameter is not 130 mm)
- once the data is added, conduct quality assurance in the form of accuracy checking the inputted data against raw data.

The resulting updated data could be eventually uploaded into the data repository (i.e. Figshare) as a new version. Version control is enabled in Figshare to record any changes to the dataset over time and to allow recalling specific versions as required.

[9] Contact Dana Clark (dana.clark@cawthron.org.nz)

10. Acknowledgments

We thank Massey University and the Ministry of Business, Innovation and Employment (MBIE) for supporting and funding this work (contract MAUX1502). We also acknowledge the support of New Zealand regional authorities that provided data and permission to use it: Northland Regional Council, Auckland Regional Council, Waikato Regional Council, Bay of Plenty Regional Council, Hawkes Bay Regional Council, Greater Wellington Regional Council, Marlborough District Council, Nelson City Council, Tasman District Council, West Coast Regional Council, Environment Canterbury, Christchurch City Council, Otago Regional Council and Environment Southland. Fiona Gower and Hugo Borges (Cawthron) and Celine Dufour (SLR Consulting NZ) provided advice on the taxonomic identification of macrofauna. Thanks to Alice Morrison for providing the cover photo for this report.

11.References

Berthelsen, A., Atalah, J., & Clark, D. (2017). *National Estuary Dataset: inconsistencies in survey data*. Prepared for Northland Regional Council. Cawthron Report No. 3107.

Berthelsen, A., Atalah, J., Clark, D., Goodwin, E., Patterson, M., & Sinner, J. (2018). Relationships between biotic indices, multiple stressor gradients and natural variability in New Zealand estuaries. *Ecological Indicators, 85*, 634-643.

Boffa Miskell. (2014). *Transmission Gully marine benthic monitoring: Report on Baseline 4 survey*. Report prepared by Boffa Miskell Limited for Wellington Gateway Partnership.

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

Chamberlain, S. *taxizesoap: Taxonomic Information from Around the Soap Web. R package version 0.1.1.9830*. https://github.com/ropensci/taxizesoap. .

Chamberlain, S., & Szocs, E. (2013). *taxize - taxonomic search and retrieval in R. F1000Research, 2:191*. http://f1000research.com/articles/2-191/v2.

Chamberlain, S., Szocs, E., Boettiger, C., Ram, K., Bartomeus, I., J., B., . . . J., O. D. (2016). *Taxonomic information from around the web. R package version 0.7.8.*: https://github.com/ropensci/taxize.

Clark, D., Berthelsen, A., Goodwin, E., Atalah, J., Patterson, M., & Sinner, J. (2018). *National Estuary Dataset*. Figshare https://doi.org/10.6084/m9.figshare.5998622.v1

Clarke, K. R., Gorley, R. N., Somerfield, P. J., & Warwick, R. M. (2014). *Change in marine communities: an approach to statistical analysis and interpretation, 3rd edition*. Plymouth: PRIMER-E

Ellis, J., Clark, D., Hewitt, J., Taiapa, C., Sinner, J., Patterson, M., . . . McCallion, A. (2013). *Ecological Survey of Tauranga Harbour*. Prepared for Manaaki Taha Moana, Manaaki Taha Moana Research Report No. 13. Cawthron Report No. 2321.

Halliday, J., Townsend, M., & Lundquist, C. (2012). *Central Waitemata Harbour ecological monitoring: 2000-2012*. Prepared by NIWA for Auckland Council. Auckland Council technical report, TR2012/027.

Hewitt, J. E., Hailes, S. F., & Greenfield, B. L. (2014). *Protocol for processing, identification and quality assurance of New Zealand marine benthic invertebrate samples*. Prepared for Northland Regional Council.

Mills, G., & Williamson, B. (2014). *Marine sediment contaminant monitoring programme: review of data quality and procedures*. Prepared by Diffuse Sources Ltd for Auckland Council. Auckland Council technical report, TR2014/041.

R Core Team. (2017). *R: A language and environment for statistical computing.* . Vienna, Austria: R Foundation for Statistical Computing https://www.R-project.org/.

Robertson, B., Gillespie, P., Asher, R., Frisk, S., Keeley, N., Hopkins, G., . . . Tuckey, B. (2002). *Estuarine Evironmental Assessment and Monitoring: A National Protocol. Part A - Development of the Monitoring Protocol for New Zealand Estuaries: introduction, rationale and methodology*. Prepared for Supporting Councils and the Ministry for the Environment, Sustainable Management Fund Contract No. 5096.

Robertson, B. M., Stevens, L., Robertson, B., Zeldis, J., Green, M., Madarasz-Smith, A., . . . Oliver, M. (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.

WoRMS Editorial Board. (2017). *World Register of Marine Species*. Available from http://www.marinespecies.org at VLIZ.

Appendix A. Option for taxonomic resolution of taxa to increase comparability of the macrofaunal data within the National Estuary Dataset

We based this option for taxonomic lumping on presence/absence shade plots (created in PRIMER 7: Clark et al. 2014), as well as on conversations with taxonomists regarding uncertainties associated with taxonomic naming and identification. Ecological differences between taxa were also considered to some extent i.e. if there were known to be important ecological differences between key taxa, every attempt was made to keep the taxa unlumped. Where there were taxa uncertainties, our general rule was to aggregate to higher taxonomic groups although in some cases it was deemed acceptable to lump to a lower taxonomic group e.g. the family Amphibolidae was lumped into the species *Amphibola crenata* because there is only one species known to belong to this family in New Zealand (Spencer et al., 2009). After consultation with taxonomic experts, lumping not based on higher taxonomic groups was conducted in two cases due to taxonomic discrepancy and uncertainty; 1) combination of two polychaete species from the Capitellidae family (*Heteromastus filiformis* and *Barantolla lepte*), and 2) combination of multiple polychaete taxa in the Spionidae family into a 'polydorid complex' grouping. Juveniles, where separately identified, were combined with parent taxa. In our analysis we removed Porifera, Tunicata, Bryozoa and Ascidiacea and these taxa are not included in the following table, however this is optional as the taxa can easily be kept in. Due to the higher-level identification of some polychaete taxa within raw data, this lumping option comes at the expense of the recommended removal of all Otago Regional Council (ORC) data collected after 2001.

The following table details the taxonomic lumping option described above aimed to increase comparability of the macrofaunal data within the National Estuary Dataset. Note that taxon names are written as they appear in the dataset, hence the lack of capital letters and italicized species names.

Reference

Spencer, H.G., Willan, R.C., Marshall, B.A., & Murray, T.J. (2009) Checklist of the recent Mollusca recorded from the New Zealand Exclusive Economic Zone. Copyright © 2016 by: Hamish G. Spencer, Richard C. Willan, Bruce A. Marshall and Tara J. Murray.
http://www.molluscs.otago.ac.nz/

Lumped taxa name	Taxa to lump
acari	acari halacaridae
actiniidae	anemone anthopleura aureoradiata
alpheidae	alpheidae alpheus alpheus socialis betaeus aequimanus
amalda	amalda amalda australis
amphibola crenata	amphibola crenata amphibola crenata juvenile amphibolidae
amphipoda	amphipoda aora maculata caprellidae caprellina longicollis corophiidae corophium dexaminidae gammaridae gammaropsis haustoriidae ischyroceridae liljeborgia liljeborgiidae lysianassidae melita awa melitidae methalimedon monocorophium monocorophium sextonae oedicerotidae paracalliope paracalliope novizealandiae paracalliopiidae paracorophium paracorophium excavatum paracorophium lucasi paradexamine paramoera chevreuxi parawaldeckia phoxocephalidae pontogeneiidae talitridae torridoharpinia torridoharpinia hurleyi urothoidae waitangi brevirostris waitangi chelatus
anthozoa	anthozoa virgularia gracillima
aonides	aonides aonides oxycephala aonides trifida

Lumped taxa name	Taxa to lump
arenicolidae	abarenicola affinis arenicolidae
armandia maculata	armandia armandia maculata
arthritica	arthritica arthritica bifurca
austrominius modestus	austrominius austrominius modestus
bivalvia	bivalvia bivalvia juvenile
calyptraeidae	sigapatella novaezelandiae zegalerus tenuis
capitella	capitella capitella capitata
chiton	chiton chiton glaucus
cirratulidae	aphelochaeta caulleriella cirratulidae
cirripedia	cirripedia sessilia
copepoda	copepoda harpacticoida
cossura consimilis	cossura cossura consimilis
crangonidae	philocheras australis pontophilus
crustacea	brachyura brachyura juvenile crustacea decapoda
cumacea	colurostylis colurostylis lemurum cumacea cyclaspis cyclaspis thomsoni diastylidae diastylopsis elongata gynodiastylis
cyclomactra	cyclomactra cyclomactra ovata
daphnia	daphnia daphnia carinata daphnia juvenile
diloma	diloma diloma aethiops diloma nigerrimum diloma subrostratum diloma zelandicum
diplodonta	diplodonta globus diplodonta zelandica
dorvilleidae	dorvillea dorvilleidae
eatoniella	eatoniella eatoniella olivacea

Lumped taxa name	Taxa to lump
edwardsia	edwardsia edwardsia leucomelos edwardsia neozelanica edwardsiidae
epitoniidae	epitoniidae epitonium tenellum
eunicidae	eunicidae eunice eunice vittata lysidice marphysa depressa marphysa disjuncta marphysa unibranchiata
gastropoda	gastropoda gastropoda juvenile
glyceridae	glycera americana glycera lamelliformis glycera lamellipodia glycera ovigera glycera russa glyceridae glyceridae juvenile hemipodia simplex
goniadidae	glycinde glycinde dorsalis glycinde trífida goniada goniada grahami goniadidae
halicarcinus	halicarcinus halicarcinus cookii halicarcinus varius halicarcinus whitei halicarcinus whitei juvenile
hemigrapsus	hemigrapsus hemigrapsus crenulatus hemigrapsus sexdentatus
hesionidae	gyptis hesionidae micropodarke oxydromus angustifrons podarkeopsis
heteromastusfiliformisandbarantollalepte	barantolla lepte heteromastus filiformis

Lumped taxa name	Taxa to lump
hiatula	hiatula hiatula nítida hiatula juvenile hiatula siliquens
holothuroidea	holothuroidea paracaudina chilensis taeniogyrus dendyi
insecta	chironomidae chironomus coleoptera collembola corynoneura scutellata dicranomyia nigrescens diptera dolichopodidae elmidae ephydridae ephydridae juvenile ephydroidea formicidae insect limnophilinae limonia microvelia muscidae orthocladiinae polypedilum stratiomyidae
isopoda	anthuridae anthuroidea cirolana woodjonesi cirolanidae eurylana eurylana arcuata eurylana cookii exosphaeroma exosphaeroma chilensis exosphaeroma falcatum exosphaeroma gigas exosphaeroma obtusum exosphaeroma planulum exosphaeroma waitemata isocladus isocladus armatus isopoda munna neozelanica munna schauinslandi munnidae natatolana natatolana pellucida paravireia paravireia pistus pseudaega melanica pseudaega punctata sphaeroma quoianum sphaeromatidae

Lumped taxa name	Taxa to lump
lasaea	lasaea hinemoa lasaea parengaensis
lumbrineridae	lumbrineridae lumbrineris scoletoma brevicirra
lunella smaragda	lunella smaragda lunella smaragda juvenile
mactridae	mactra mactra ordinaria mactridae
magelona	magelona magelona dakini magelona papillicornis
maldanidae	asychis axiothella serrata euclymene macroclymenella stewartensis maldanidae maldanidae juvenile
micrelenchus	micrelenchus micrelenchus huttonii micrelenchus tenebrosus
microspio	microspio microspio maori
mysella	mysella mysella juvenile
mysida	Mysida mysidae tenagomysis
mytilidae	musculus impactus mytilidae mytilidae juvenile mytilus mytilus edulis mytilus galloprovincialis mytilus juvenile perna canaliculus xenostrobus pulex
nebaliacea	nebalia nebaliacea
nemertea	adenorhagas aurantiafrons nemertea
neoguraleus	neoguraleus neoguraleus sinclairi
nephtyidae	aglaophamus aglaophamus macroura nephtyidae
nereididae	ceratonereis neanthes neanthes cricognatha nereididae nereididae juvenile nereis nereis falcaria perinereis perinereis brevicirris

Lumped taxa name	Taxa to lump
	perinereis camiguinoides perinereis nuntia brevicirris perinereis vallata platynereis platynereis australis
notoacmea	notoacmea notoacmea elongata notoacmea scapha
notomastus	capitellethus zeylanicus notomastus notomastus zeylanicus
nuculidae	linucula hartvigiana nucula nucula gallinacea nucula nitidula
onuphidae	diopatra akarana onuphidae onuphis aucklandensis
ophiuroidea	amphiura ophionereididae ophiurida ophiuroidea
opisthobranchia	aglajidae bulla quoyii melanochlamys cylindrica nudibranchia nudibranchus opisthobranchia philine philine auriformis relichna aupouria
orbiniidae	leitoscoloplos leitoscoloplos kerguelensis naineris naineris grubei australis phylo novaezealandiae orbiniidae
ostracoda	copytus novaezealandiae cypridinodes concentrica cypridinodes reticulata cytherella diasterope grisea euphilomedes agilis leuroleberis zealandica ostracoda parasterope parasterope quadrata rutiderma
ostreidae	crassostrea gigas ostrea chilensis ostreidae juvenile saccostrea cucullata glomerata
owenia	owenia fusiformis owenia petersenae
paguridae	paguridae paguristes pagurus

Lumped taxa name	Taxa to lump
palaemonidae	palaemon palaemon affinis palaemonidae
paraonidae	aricidea levinsenia gracilis paradoneis paradoneis lyra paraonidae
paraprionospio	paraprionospio paraprionospio coora
pectinariidae	pectinaria pectinariidae
phyllodocidae	eteone eulalia microphylla phyllodocidae
platyhelminthes	platyhelminthes stylochidae
polychaeta	phyllodocida polychaeta
polydorid complex	boccardia boccardia acus boccardia juvenile boccardia knoxi boccardia polybranchia boccardia syrtis polydora polydora cornuta pseudopolydora pseudopolydora paucibranchiata
polynoidae	antinoe disconatis accolus frennia harmothoe lepidastheniella comma lepidonotinae lepidonotus lepidonotus polychromus paralepidonotus ampulliferus polynoidae polynoinae
potamopyrgus	potamopyrgus potamopyrgus antipodarum potamopyrgus estuarinus
prionospio	prionospio prionospio aucklandica prionospio cirrifera prionospio ehlersi prionospio yuriel
pycnogonida	pantopoda pycnogonida pycnogonidae
sabellidae	euchone euchone pallida pseudopotamilla sabellidae

Lumped taxa name	Taxa to lump
scalibregmatidae	hyboscolex longiseta scalibregma inflatum scalibregmatidae
scolecolepides	scolecolepides scolecolepides benhami
serpulidae	serpulidae spirobranchus spirobranchus cariniferus
sigalionidae	labiosthenolepis laevis sigalionidae
sipuncula	sipuncula sipunculidae
sphaerodoridae	sphaerodoridae sphaerodoropsis
spionidae	pseudonerine rhynchospio spio spionidae spiophanes kroyeri
stomatopoda	heterosquilla lysiosquilla squillidae stomatopoda
syllidae	exogone exogoninae sphaerosyllis sphaerosyllis hirsuta sphaerosyllis semiverrucosa syllidae syllinae syllis
terebellidae	streblosoma toddae terebellidae terebellinae
travisia	travisia travisia olens travisia olens novaezealandiae
trichobranchidae	terebellides stroemii trichobranchidae
trichoptera	trichoptera rhyacophiloidea
venerida	irus reflexus ruditapes largillierti juvenile venerida
xymene	xymene xymene ambiguus xymene plebeius
zeacumantus	zeacumantus zeacumantus lutulentus zeacumantus subcarinatus

The taxa to remain as they are (i.e. unlumped) are:

acanthochitona zelandica, ampharetidae, annelida, antisolarium egenum, araneae, arcuatula senhousia, asteroidea, austrofusus glans, austrohelice crassa, austrovenus stutchburyi, biffarius filholi, borniola reniformis, buccinulum, capitellidae, chaetognatha, cidaridae juvenile, cominella adspersa, cominella glandiformis, cominella maculosa, corbula zelandica, cyclograpsus lavauxi, divalucina cumingi, dosinia subrosea, enteropneusta, euterebra tristis, fellaster zelandiae, flabelligeridae, halopyrgus pupoides, haminoea zelandiae, haustrum scobina, hemiplax hirtipes, hirudinea, hunkydora australica, hydrozoa, ischnochiton maorianus, leptomya retiaria retiaria, macomona liliana, manayunkia, melanopsis, melliteryx parva, myadora, myllitella vivens vivens, nassarius burchardi, nematoda, neosabellaria kaiparaensis, nepinnotheres atrinicola, nepinnotheres novaezelandiae, nerita melanotragus, nicon aestuariensis, odostomia, oenonidae, oligochaeta, opheliidae, orbinia papillosa, ovalipes catharus, oweniidae, paphies australis, paphies donacina, paratya curvirostris, patiriella regularis, peronaea gaimardi, perrierina turneri, phoronida, phyllochaetopterus socialis, pisinna zosterophila, pradoxa, pseudarcopagia, rhyssoplax, risellopsis varia, rissoidae, scolelepis, scoloplos, scoloplos cylindrifer, solemya parkinsonii, sypharochiton pelliserpentis, tanaidacea, theora lubrica, trochus tiaratus, turbonilla, turridae, zalipais lissa, zethalia zelandica.

Appendix B. Detailed examples of laboratory analysis methods for sediment physico/chemical variables in the National Estuary Dataset

Note that this is not necessarily an exhaustive list.

Variable	**Laboratory analysis [information source]**
Sediment grain size - laser	Sediments were pre-treated with 10% hydrogen peroxide to remove organic material and 1M hydrochloric acid to remove carbonate material. Calgon™ was added as a dispersant and samples were placed in an ultrasonic bath for 10 minutes to aid disaggregation. Samples were analysed using a Malvern Mastersizer 2000. Grain size data were grouped into the following grain size categories: mud (<63 µm); very fine sand (63-125 µm); fine sand (125-250 µm); medium sand (250-500 µm); coarse sand (500-1000 µm) and gravel (>1000 um) (following the Wentworth sediment classification). [Needham et al. 2014] (Report only until 2011 but assume the same analysis used from 2012 onwards.) Samples were analysed by Auckland University Services Ltd with a laser diffraction particle analyser (Malvern Mastersizer 2000). The following size fractions were determined: < 63 µm (mud); 63 -230 µm (fine sand); 250-500 µm (medium sand); and >500 µm (coarse sand). [Griffiths 2011]
Sediment grain size – wet sieve	Wet sieving, gravimetry (calculation by difference) [Hill Laboratories Analysis Report Quote 31586 GWRC Porirua 2008]. Sieving, gravimetric. All drying 35 °C, overnight [Hill Laboratories Analysis Report Quote 439846 ORC Waikauaiti 2006, Smith 2009]. In House Method [Cawthron Laboratory Report number S84798 Tauranga 2011, Madarasz 2006] Wet sieving and calculation of percentage fractions according to dry weight [Robertson et al. 2002, Gillespie & Clark 2007]

Variable	Laboratory analysis [information source]
	< 63 µm Wet Sieved with no gravimetric determination. [Boffa Miskell Limited 2014][10] The samples are homogenised and a subsample of approximately 5 g of sediment taken, and digested in ~ 9% hydrogen peroxide until frothing ceases. The sediment sample is then wet sieved through 2000 µm, 500 µm, 250 µm and 63 µm mesh sieves. Pipette analysis is used to separate the <63 µm fraction into >3.9 µm and <3.9 µm. All fractions are then dried at 60°C until a constant weight is achieved (fractions are weighed at ~ 40 h and then again at 48 h). The results of the analysis are presented as percentage weight of gravel/shell hash (>2000 µm), coarse sand (500 – 2000 µm), medium sand (250 – 500 µm), fine sand (62.5 – 250 µm), silt (3.9 – 62.5 µm) and clay (<3.9 µm). [Halliday et al. 2012] Prior to analysis, the samples are homogenised and a subsample of approximately 5 g of sediment taken. They are then digested in 6% hydrogen peroxide until all organic matter is removed, and sampled by wet sieving and pipette analysis (Gatehouse 1971). Pipette analysis is used to separate the <63 µm fraction into >3.9 µm and <3.9 µm. All fractions are then dried at 60°C until a constant weight is achieved (fractions are weighed at ~ 40 hr and then again at 48 hr). The results of the grain size analyses are presented as percentage composition of gravel/shell hash (>2 mm), coarse sand (500–2000 µm), medium sand (250–500 µm), fine sand (62.5–500 µm), silt (3.9–62.5 µm) and clay (<3.9 µm). Mud content is calculated as the sum of the silt and clay content. [Greenfield et al. 2016] Prior to grainsize analysis, organic matter was removed using 9% hydrogen peroxide until fizzing ceased. Samples were then dried and weighed to obtain a total dry weight. They were then deflocculated for at least 4 hours (using Calgon 5 g per litre) and wet-sieved on a stack of sieves (500, 250, 125 and 63 µm). Each fraction was dried, weighed and calculated as a percentage of the total weight. The fraction less than 63 µm was calculated by subtraction of all other dry weights from the initial dry weight. Sediment % weight was then expressed for coarse sand (> 500), medium sand (250–499), fine sand (125–249), very fine sand (63–124) and mud (< 63 µm). Sampling in Whangateau initially used the sampling protocol in the ecological monitoring programmes conducted in Manukau, Mahurangi and Central and Upper Waitemata Harbours. In these programmes, very fine sand and fine sand were not separated, but three additional fractions were calculated: % gravel (>2 mm); and the mud component was separated by pipette analysis into % silt (4 – 63 m) and % clay (<3.9 m). However, from 2011, samples have been analysed as above. [Hewitt & Simpson 2012]
Metals	Dry weight by ICPMS – USEPA 200.8 (Modified) [Watercare Laboratory Sampling Number MON-005477, Kerikeri, 2008 NRC]

[10] Consent monitoring data. This was assigned the wet sieving methodology in the dataset, even though laser analysis is also mentioned, as grains > 2000 µm were analysed.

Variable	Laboratory analysis [information source]
	Dried sample, sieved as specified (if required). Nitric/Hydrochloric acid digestion, ICP-MS, trace level. US EPA 200.2. [Hill Laboratories Report Number 627385 Porirua GWRC 2008] Nitric / hydrochloric acid digestion, ICP-MS (Low level). US EPA 200.2 [Hill Laboratories Report Number 439846 Waikauaiti 2006, Madarasz 2006] Dried sample, <2mm fraction. Nitric/Hydrochloric acid digestion, ICP-MS, screen level. US EPA 200.2 [Hill Laboratories Report Number 618099 Kaikorai 2007] Dried sample, <2mm fraction. Nitric/Hydrochloric acid digestion, ICP-MS, trace level. US EPA 200.2 [Hill Laboratories Report Number 618099 Kaikorai 2007] Dry/sieve sample, Digestion US EPA 200.2. Air dry 35°C/2mm sieve Nitric/HCl acid digestion, ICP-MS [Smith 2009] Dried sample, <2mm fraction. Nitric/Hydrochloric acid digestion, ICP-MS, trace level. [Hill Laboratories Report Number 1248339 Waimea 2014] USEPA 200.2 Digestion / ICP-MS [Cawthron Laboratory Report Number S84798 Tauranga 2011] Perchloric/nitric acid digestion and flame atomic absorption spectrometry (ASTM 3974 Digestion Practice A; AOAC 1995 950.46 modified) [Robertson et al. 2002, Gillespie & Clark 2007] Chemical analysis was performed on total recoverable acid digested < 500 µm dry sieved fractions for all metals. [Hewitt & Simpson 2012]
TOC	Acid pre-treatment to remove carbonates if present, neutralisation, Elementar Combustion Analyser. [Boffa Miskell Limited 2014] Acid pre-treatment to remove carbonates if present, Elementar Combustion Analyser. [Hill Laboratories Report Number 1248339 Waimea 2014, Hill Laboratories Report Number 1401330 Havelock, 2015]

Variable	Laboratory analysis [information source]
	Sediments were dried and finely ground, then analysed for total organic carbon content using an automated CHN analyser. Samples were pre-treated with acid to remove carbonate material prior to analysis [Needham et al. 2014] (Report only until 2011 but assume the same analysis used from 2012 onwards.)
AFDW	Ignition in muffle furnace 550°C, 6hr, gravimetric. APHA 2540 G 21st ed. 2005. [Hill Laboratories Report Number 627385 Porirua GWRC 2008] Ignition in muffle furnace 550°C, 1hr, gravimetric. (Also called Volatile Matter or Ash Free Dry Weight) APHA 2540 G 20th ed. 1998 [Hill Laboratories Report Number 439846 Waikauaiti 2006, Madarasz 2006] APHA 21st Edn 2540 D+ E (Mod) [Cawthron Laboratory Report number S84798 Tauranga 2011, Smith 2009] APHA 20th Edn 2540D+ E (Mod) [Madarasz 2006] Weight loss from dry sediment after combustion at 550°C (APHA 1999, 20th Edn, modified 2540D + E) [Robertson et al. 2002] Approximately 5 g of sediment is placed in a dry, pre-weighed tray. The sample is then dried at 60°C until a constant weight is achieved (the sample is weighed after ~ 40 h and then again after 48 h). The sample is then ashed for 5.5 h at 400°C (Mook and Hoskin 1982) and then reweighed. [Halliday et al. 2012]
TP	Dry Weight by ICP-MS – USEPA 200.8 (Modified) [Watercare Laboratory Sampling Number MON-005477 Kerikeri NRC 2008] Dried sample, sieved as specified (if required). Nitric/Hydrochloric acid digestion, ICP-MS, screen level. USEPA 200.2. [Hill Laboratories Report Number 627385 Porirua GWRC 2008] Nitric / hydrochloric acid digestion, ICP-MS. US EPA 200.2 [Hill Laboratories Report Number 439846 Waikauaiti 2006, Madarasz 2006] Dried sample, <2mm fraction. Nitric/Hydrochloric acid digestion, ICP-MS, screen level. US EPA 200.2 [Hill Laboratories Report Number 618099 Kaikorai 2007]

Variable	Laboratory analysis [information source]
	ICP-MS Aqua Regia Digest [Gillespie & Clark 2007] USEPA 200.2 Digestion / ICP-MS [Cawthron Laboratory Report Number S84798, Tauranga 2011] Colourimetric (APHA, 20th Edn. 1999, Method 4500-P. A, B, E) [Robertson et al. 2002]
TKN	Distillation, colourimetric (APHA, 19th Edn. 1995, Method 4500-N Org C) [Robertson et al. 2002]
TN	IN HOUSE [Watercare Laboratory Sampling Number MON-005477 Kerikeri NRC 2008] Catalytic Combustion (900°C, O2), separation, Thermal Conductivity Detector [Elementar Analyser]. [Hill Laboratories Report Number 627385 Porirua GWRC 2008, Smith 2009, Madarasz 2006] Catalytic Combustion, separation, Thermal Conductivity Detector [Elementar Analyser]. [Hill Laboratories Report Number 1248339 Waimea 2014] APHA 21st Edn 4500N C [Cawthron Laboratory Report Number S84798 Tauranga 2011] Sediments were dried and finely ground, then analysed for total nitrogen content using an automated CHN analyser [Needham et al. 2014] (Report only until 2011 but assume the same analysis used from 2012 onwards.) APHA 20th Edn 4500N C [Gillespie & Clark 2007]

Additional references

Gillespie, P., Clark, K. (2007). *Orowaiti estuary fine-scale benthic baseline 2007*. Prepared for West Coast Regional Council. Cawthron Report No. 1274. 19p plus appendices.

Hewitt, J., Simpson, J. (2012). *Assessment of the estuarine ecological monitoring programme to 2012*. Prepared by NIWA for Auckland Council. Auckland Council technical report 2012/040.

Gatehouse, J.S. (1971). Sedimentary analysis. In Carver, R.E. (Ed.), *Procedures in Sedimentology and Petrology*. Wiley Interscience, New York, pp. 59-94.

Greenfield, B.L., Hewitt, J.E., Hailes, S.F. (2016). *Manukau Harbour ecological monitoring programme: report on data collected up until February 2015*. Prepared by the National Institute for Water and Atmospheric Research, NIWA for Auckland Council. Auckland Council technical report, TR2016/029.

Griffiths, R. (2011). *Ruakaka estuary: estuary monitoring programme results from 2008 – 2010*. Northland Regional Council.

Madarasz, A. (2006). Ahuriri Estuary: Environmental Assessment and Monitoring. Internal Environmental Monitoring.

Mook, D.H., Hoskin, C.M. (1982). Organic determination by ignition: caution advised. *Estuarine Coastal and Shelf Science*, *15*, 697-699.

Needham, H., Singleton, N., Giles, H., Jones, H. (2014). *Regional Estuary Monitoring Programme 10 year trend report: April 2001 to April 2011*. Waikato Reginal Council Technical Report 2014/41.

Smith, S. (2009). *Estuarine ecology programme: environmental assessment of Ahuriri and Porangahau estuaries*. Prepared for Hawke's Bay Regional Council by EAM Environmental Consultants.