Suitability of a New Functional Traits Index as a State of the Environment Indicator

Technical Report: 2011/004

Auckland Council
Technical Report 2011/004
ISSN 2230-4525 (Print)
ISSN 2230-4533 (Online)

ISBN 978-1-927169-06-3 (Print)
ISBN 978-1-927169-07-0 (PDF)

Reviewed by:

Name: Megan Carbines

Position: Scientist – Marine

Organisation: Auckland Council

Date: 15.04.2011

Approved for Auckland Council Publication by:

Name: Grant Barnes

Position: Manager Research Investigation and Monitoring

Organisation: Auckland Council

Date: 30.04.2011

Recommended Citation:
Lohrer, D.; Rodil, IF. (2011). Suitability of a New Functional Traits Index as a State of the Environment Indicator. Prepared by NIWA for Auckland Council. Auckland Council Technical Report 2011/004

Suitability of a New Functional Traits Index as a State of the Environment Indicator

Drew Lohrer
Ivan Franco Rodil

Prepared for
Auckland Council

National Institute of Water & Atmospheric Research Ltd
Gate 10 Silverdale Road
Hillcrest, Hamilton 3216

Client Report No: HAM2011-013

Contents

1	**Introduction**	**1**
1.1	Background	1
1.2	Objectives	1
1.3	Review of the original NIWACOOBII methodology	2
2	**Methods**	**5**
2.1	Revision of the functional traits database	5
2.2	The SUMmax parameter	6
2.2.1	Accounting for richer sites	6
2.2.2	Accounting for differences in numbers of replicates	6
2.3	Data for NIWACOOBII testing	7
2.4	NIWACOOBII calculations	9
2.5	Taxonomic resolution and NIWACOOBII scores	9
3	**Results**	**10**
3.1	The functional traits database	10
3.2	SUMmax$_n$	11
3.3	NIWACOOBII calculations	17
3.4	Taxonomic resolution and NIWACOOBII scores	24
4	**Discussion**	**26**
4.1	NIWACOOBII performance	27
4.2	Conclusions and recommendations	28
5	**Acknowledgements**	**30**
6	**References**	**31**

Reviewed by:

Dr Simon F. Thrush

Approved for release by:

Dr Judi E Hewitt

Formatting checked

Executive Summary

The Auckland Council requires easily understandable but scientifically defensible indicators of the ecological integrity of its estuarine and coastal areas. A functional indicator developed in New Zealand in 2010 with local knowledge, called the NIWACOOBII, was responsive to strong gradients of mud and heavy metal contaminants in the sediments of Auckland area estuaries. The NIWACOOBII performed better than available indices that had been developed overseas and showed promise as a potential ecological indicator for use in State of the Environment reporting. However, further testing and validation of the NIWACOOBII with independently collected data was deemed necessary prior to adoption and use by the Auckland Council.

The NIWACOOBII index is based upon the richness of macrofaunal taxa in seven individual functional trait groups. The index tracks a broad cross section of macrofaunal functional types, with one trait group selected from each of seven broader functional trait categories (organism size, mobility, feeding mode, position in the sediment, etc.). The seven individual trait groups selected for use in the index were those most sensitive to mud and metals. The index runs from 0 to 1, with values near 0 indicating highly degraded sites and values near 1 indicating the opposite. Declines in NIWACOOBII scores with increases in mud and heavy metals are interpreted as losses of functional redundancy. Habitats with high functional redundancy (i.e., many species present in each functional trait group) will tend to have higher inherent resistance and resilience in the face of environmental changes, as the higher numbers of species per functional group provide “insurance” for stochastic or stress-induced losses of particular species.

Adjustments have been made to the way in which the NIWACOOBII is calculated after trialling the index at monitoring sites in three harbours and identifying ways in which it could be improved. Specifically, the functional traits database that underpins the NIWACOOBII was updated in light of new knowledge and the appearance of new species in our data sets. Secondly, a parameter involved in the NIWACOOBII calculation called SUMmax was adjusted to accommodate high observed SUMactual values. SUMmax also now accounts for the effect of sample size on SUMactual scores. By making these adjustments, the NIWACOOBII calculations are more robust and can be validly compared across new sites and sites sampled by differing numbers of replicates.

Once the calculation method was settled, we tested the NIWACOOBII using mud, heavy metal and macrofaunal data from 34 sites. None of the data from these 34 sites was used during the initial development phase of the NIWACOOBII, ensuring a truly independent test. NIWACOOBII scores were significantly lower at sites with high levels of mud and heavy metals. Decreases in NIWACOOBII scores with increasing heavy metal concentrations were observed for Cu, Pb and Zn individually and for the metal gradient indices CCU and PCA1.500. However, the percent of variation in NIWACOOBII scores explained by mud or heavy metals was rather low (<25%).

Although the NIWACOOBI is less sensitive to mud and heavy metal pollution gradients than the previously developed Benthic Health Model, the NIWACOOBII provides more information on whether functional redundancy is changing and whether specific functional traits are being affected. Furthermore, the NIWACOOBII can be validly calculated in places with different regional species pools (for only the presence of particular functional traits is tallied, rather than particular species), whereas the BHM is regionally restricted. Therefore, for use in State of Environment reporting, the NIWACOOBII and BHM may complement each other well by providing a balance of sensitivity, information content and broad general applicability.

1 Introduction

1.1 Background

The Auckland Council requires easily understandable but scientifically defensible indicators of the ecological integrity of its estuarine and coastal areas. These indicators can facilitate non-technical communication in State of Environment reporting or can be used more generally for informing the public about the health status of highly valued coastal habitats.

Overseas, a number of indices of ecological integrity have been developed, but there has been considerable scientific debate about their interpretation and validity. In 2009-2010, NIWA was commissioned to review some of the available overseas indicators and determine whether or not they could be adapted for use in the Auckland region. Two indices, AMBI (Borja et al. 2008) and B-IBI (Weisberg et al. 1997), were applied to Auckland Regional Council data sets to determine how well they correlated with gradients of heavy metal concentration and sediment mud content in Auckland area estuaries (van Houte-Howes and Lohrer 2010).

The two overseas indicators (AMBI, B-IBI) did not effectively track stress gradients in Auckland Region intertidal habitats and were difficult to use objectively. A functional indicator developed in New Zealand with local knowledge, the NIWACOOBII, was more effective (van Houte-Howes and Lohrer 2010). Further testing and validation of the NIWACOOBII with independently collected data was deemed necessary prior to adoption and use by the Auckland Council for State of Environment reporting.

It was important that the NIWACOOBII effectively track gradients of mud and heavy metals, as these are the two predominant pollutant types affecting intertidal flats in the Auckland area. However, the univariate NIWACOOBII was not expected to have the same degree of sensitivity and explanatory power as the multivariate Benthic Health Model (Anderson et al. 2002, 2006). NIWACOOBII scores are correlated with BHM scores (van Houte-Howes and Lohrer 2010), although the NIWACOOBII appears to provide information and apply to areas that the BHM does not. Thus, the ability of the NIWACOOII to complement the BHM was considered as part of the NIWACOOBII evaluation.

1.2 Objectives

The overarching aim of the work reported here was to test the functional index called the NIWACOOBII and to report on its applicability for SOE reporting. However, prior to index testing, we sought to improve on the NIWACOOBII calculations as presented by van Houte-Howes and Lohrer (2010). The modifications were designed to facilitate the calculation of NIWACOOBII values in various contexts (i.e., monitoring data sets) and to ensure the comparability of NIWACOOBII calculations from this point forward. Specifically,

- Some species were added to the present NIWA functional traits database (which is the basis of the NIWACOOBII) and the database was updated in light of our growing knowledge of species traits. For each taxon (~480 in all), scores for 32 corresponding functional trait categories were assigned.
- A key parameter in the NIWACOOBII calculation, called SUMmax, needed to be re-set. This was necessary for two reasons:
  - The SUMmax value used to standardize the original NIWACOOBII calculation (van Houte-Howes and Lohrer 2010) was found to be too low for two Central Waitemata Harbour sites (Reef and Shoal Bay, see Townsend 2010) and was therefore likely to be too low for other sites as well.
  - It was likely that sample size (number of macrofaunal cores collected per site) would have an effect on SUMmax values, and this issue needed to be resolved.

Once the basic NIWACOOBII calculations had been improved, we tested the ability of the index to track gradients of mud and heavy metals at an independent set of field sites. Specifically, as the NIWACOOBII was developed using metal and mud content data from 95 Regional Discharges Project and 5 Mahurangi Harbour sites (van Houte-Howes and Lohrer 2010), it needed to be validated using data from new sites. We compiled mud, heavy metal and macrofauna community data from 34 new sites and used these data for index testing.

Our final objective was to test the influence of taxonomic resolution (identifications to family level only versus to genus and species level) on NIWACOOBII index values and to make recommendations on the need for identifications to species level.

1.3 Review of the original NIWACOOBII methodology

The NIWACOOBII was initially developed using a master database of coastal soft-sediment taxa found in New Zealand with their associated functional traits (Hewitt et al. 2008). There were 7 broad categories of functional traits, with 29 functional trait groups in total (van Houte-Howes and Lohrer 2010)[1]. The assignment of functional traits to taxa was done using in-house knowledge and the best available information from the literature. The functional groupings were based on macrofaunal attributes that included feeding behaviour, position in the sediment column, degree of motility, type of topographic feature created (tubes/pits/mounds), body size, body shape, and so on. When the role of a taxon did not fall distinctly into one category or another, or when little information was available about traits, “fuzzy” probabilities were used. For example, some taxa live throughout the sediment column and are thus found near the sediment surface and down deeper; these were coded as both “Top” (found in the upper 0-2 cm of the sediment) and “Deep” (found in the 2-10 cm sediment horizon) by assigning each code a value of 0.5. As another example, when we had little or no information on mobility traits for a taxon, we coded the taxon as having equal probabilities (0.33) of being sedentary, limited mobility or highly mobile.

[1] This has been modified slightly since publication of van Houte-Howes and Lohrer (2010). There are now 8 broad categories and 32 functional trait groups (see Methods section).

Only a small subset of the species from the master database is present at any particular site and time. During NIWACOOBII development, the number of taxa in each of the 29 functional groups ($Ntaxa_{group}$) was calculated at 100 intertidal estuarine sites in the Auckland Region. Sediment mud content and heavy metal contaminant data were available for all sites (van Houte-Howes and Lohrer 2010).

Using this data set, correlations between stress levels (mud percentage and heavy metal concentration) and functional composition ($Ntaxa_{group}$) were calculated. We tabulated results on the number of correlations that were positive versus negative (i.e., increasing versus decreasing $Ntaxa_{group}$ with increasing stress) and also examined the strength of the correlation coefficients (Pearson's *r* values). The analysis of the 29 functional groupings revealed consistent responses to increasing stress levels for both mud and metals. Generally, the number of taxa present per group, $Ntaxa_{group}$, decreased with increasing stress levels (27 negative correlations with mud, 27 negative correlations with metals, no correlations were significantly positive) (Tables 6 and 7 of van Houte-Howes and Lohrer 2010).

Seven of the original 29 functional groups were retained for use in the index, with one grouping selected from each broader functional category (Table 1). All seven of the groups included had strong and significant negative responses to both mud and metals (average $r < -0.5$). The seven selected groups were "Top 2 cm" (organisms that occupy the upper 2 cm of the sediment column), "Erect" (organisms that create erect topographic features, such as tubes, that stick out of the sediment), "Surface-to-Surface" (organisms whose activities move sediment particles laterally across the sediment surface, as opposed to up or down), "Sedentary" (organisms that do not move, or only do so within a fixed tube), "Suspension feeders" (organisms that feed by filtering suspended particles from seawater), "Medium" (organisms of intermediate body size), and "Worm" (worm-shaped organisms with length much greater than width).

Index values were then calculated as follows:

1. The 7 selected $Ntaxa_{group}$ values per site were summed (i.e., $Ntaxa_{Top}$ + $Ntaxa_{Erect}$ + $Ntaxa_{SS}$ + $Ntaxa_{Sedentary}$ + $Ntaxa_{Sus}$ + $Ntaxa_{Medium}$+ $Ntaxa_{Worm}$) to produce a quantity called SUM_{actual}. These sums were calculated for all 100 sites.

2. A maximum expected value (i.e., a non-polluted reference value) was determined from the sums of maximum values observed across all the sites, e.g., $Ntaxa_{TopMAX}$ + $Ntaxa_{ErectMAX}$ + $Ntaxa_{SSMAX}$ + $Ntaxa_{SedentaryMAX}$ + $Ntaxa_{SusMAX}$ + $Ntaxa_{MediumMAX}$+ $Ntaxa_{WormMAX}$. The quantity was called SUM_{max}. In the original NIWACOOBII calculations (van Houte-Howes and Lohrer 2010), SUMmax was a constant with a value of 144.

3. A minimum possible value (i.e., a completely defaunated site) was set at 0.

4. The NIWACOOBII formula was $1 – (SUM_{max} – SUM_{actual})/ SUM_{max}$, which essentially standardised the index values to fall between 0 and 1. Values near 0 would indicate highly degraded sites, and values near 1 would indicate the opposite.

Table 1:

A listing of eight broad functional trait categories (left column) and the 32 individual functional trait groups among them (centre column). Asterisks next to a functional category name indicate that fuzzy probabilities were used to assign values to the corresponding trait groups during the development of the master database (see text for explanation of fuzzy probabilities). The NIWACOOBII index is based on seven of the individual trait groups, which are highlighted in grey. The eighth category, body hardness, was added in 2011; none of the body hardness trait groups factor into the NIWACOOBII calculation (see methods section 2.1).

Functional Category	Functional Group	Code
Living position *	Attached	Attached
	Deeper than 2 cm	Deep
	Surface epifauna	Epif
	Top 2 cm	Top
Sediment topography	Permanent burrow	Burr
feature created *	Erect structure / tube	Erect
	Simple hole or pit	Hole
	Mound	Mound
	Trample marks	Trample
	Trough	Trough
Direction of sediment	Depth to depth	DD
particle movement *	Depth to surface	DS
	Surface to depth	SD
	Surface to surface	SS
Degree of motility	Freely motile on or in sediment	Free
	Limited movement, usually in sediment	Limited
	Sedentary / movement in a fixed tube	Sedentar
	Semi-pelagic	Spel
Feeding behaviour *	Deposit feeder	Dep
	Grazer	Grazer
	Predator	Pred
	Scavenger	Scav
	Suspension feeder	Sus
Body size	Large	Large
	Medium	Medium
	Small	Small
Body shape	Streamlined (length 3-10x width)	Streamlined
	Round/Globulose (length 1-3x width)	Globular
	Worm-shaped (length 10-100x width)	Worm
Body hardness	Soft-bodied	Soft
	Rigid (chitonous endo- or exo-skeleton)	Rigid
	Calcified (fully calcified shell; molluscs)	Calcified

2 Methods

2.1 Revision of the functional traits database

An extensive database of species and corresponding functional traits information is required for NIWACOOBII calculations. This database will need to be updated occasionally because new species are likely to be encountered when new sites and times are sampled and because our knowledge of the natural history of species (and thus their functional traits) will continue to improve. The possibility of new arrivals of non-indigenous species also exists.

As we compiled the macrofauna data from new sites (466 macrofaunal core samples in total), several taxa needed to be added to the database. Although most of the species added were previously described species from New Zealand, there was at least one non-indigenous species that required inclusion (*Nassarius burchardi*, Townsend et al. 2010).

It was necessary to add categories such as “amphipod (unspecified)” and “gastropod (unspecified)” for specimens that were noted as being damaged or otherwise unidentifiable. It is not uncommon to find a species (e.g., *Orbinia papillosa*) and a con-familial (“orbiniid”) on the same data sheet. In these cases, we assumed that the unspecified organism (“orbiniid”) was different from the identified species (“*Orbinia papillosa*”) and thus they were listed separately in the database. The functional traits assigned to taxonomically related listings that were “unspecified” were given equal probabilities of belonging to the range exhibited by the broad taxonomic group.

As the functional traits database was updated, all of the taxonomic names were checked using the World Register of Marine Species (http://www.marinespecies.org). This was to ensure consistency in the placement of species and genera into higher-order taxonomic groupings (family, order, subclass, phylum). For the species that have undergone name revisions recently (e.g., *Aquilaspio aucklandica/Prionospio aucklandica, Helice crassa/Austrohelice crassa*), the former and current names were listed next to each other to avoid confusion.

The original functional traits database had 7 broad categories of traits and 29 individual trait groups. One of the broad categories called “Body Shape/Type” contained three trait groups called “worm-shaped”, “globular-shaped” and “calcium-shelled” (van Houte-Howes and Lohrer 2010). In our revision of the functional traits database, we changed the three groups within the Body Shape category to “worm-like”, “streamlined” and “globular/round”. Worm-like organisms were defined as animals whose length was 10 to 100 times greater than width (e.g., most polychaetes). Streamlined organisms were defined as animals with bilateral symmetry whose length was 3 to 10 times greater than width (e.g., isopods and amphipods). Globular/round organisms were those that were amorphous, spherical or approximately disk-shaped and whose body length was ~1 to 3 times the width. We then added a new category called “Body Hardness” with three groups: “soft-bodied”, “rigid” and “calcified”. Organisms in the soft-bodied group were those lacking hard parts, with the exception of jaws and setae (e.g., polychaetes, nemerteans, sipunculids, anemones). Taxa in the “rigid” group were those with

chitonous exoskeletons (e.g., crustaceans) and those with internal calcium carbonate tests/skeletons (e.g., most echinoderms). The "calcified" group was reserved mainly for gastropod and bivalve mollusks with well developed calcium carbonate shells. The "Body Hardness" category is likely to be useful for identifying the sensitivity of taxa to physical disturbance (i.e., dredging) and ocean acidification, but is unlikely to be responsive to mud or heavy metal contamination. For this reason, it was not considered necessary to revise the NIWACOOBII to include a group from this category.

Although the alteration of trait categories and groupings improves the functional traits database for a variety of future uses, the changes did not markedly affect the NIWACOOBII. This is because the classifications of taxa into the "Worm-Shaped" group (one of the seven groups used to calculate the index) was essentially unaffected.

2.2 The SUMmax parameter

2.2.1 Accounting for richer sites

The NIWACOOBII calculation is an index based on a quantity called SUMactual, which is generated from the observed taxonomic richness within seven functional trait groups in a set of samples (refer to section 1.3 above).

The NIWACOOBII also relies on the definition of a theoretical maximum value, called SUMmax, which represents the healthiest possible non-polluted reference site and is used to standardize the index to run between 0 and 1. The SUMmax value needs to be set well above the magnitude of observed SUMactual values and potential future SUMactual values as well.

The SUMmax value set during NIWACOOBII development was 144 (van Houte-Howes and Lohrer 2010). This SUMmax value was greater than any of the original SUMactual values (from 100 sites) and was therefore thought to be appropriate. SUMactual values were subsequently calculated for sites in the Kaipara, Whangateau and Central Waitemata Harbours (a total of 18 new sites). At two of the sites (Reef and Shoal Bay from CWH), SUMactual values exceeded the theoretical SUMmax, resulting in NIWACOOBII scores >1.0 (Townsend 2010). Both of these sites were highly heterogeneous in habitat. The original SUMmax value of 144 was thus too low and required a re-evaluation.

2.2.2 Accounting for differences in numbers of replicates

Many of the Auckland Council monitoring programmes are based upon 12 replicates per site, a level of replication that was determined from analyses of sampling precision (e.g., the amount of change in the variance with increasing replication)(Thrush et al. 1988). However, many other sites have been sampled with 10 replicates (i.e., RDP sites) or 9 or 6 or even 3 replicates (depending on the research question and purpose of the sampling). SUMactual values from sites with differing numbers of replicates are likely to vary with sampling intensity (irrespective of pollutant levels) because taxonomic richness invariably increases with the number of samples collected per site. We attempted to reduce this bias by specifying replicate-specific SUMmax values,

based on an analysis of species accumulation curves, for the new and improved NIWACOOBII calculations. [2]

To set SUMmax values for NIWACOOBII calculations based on 3, 6, 9, 10, and 12 replicate core samples, we examined species accumulation curves at four species rich sites (Waiheke Island, Pollen Island, Reef and Shoal Bay). Forty-eight replicate cores had previously been collected at Pollen Island and Waiheke Island (March 2006) during FRST funded research. These sites were shelly-sandy sites and both were located inside Department of Conservation marine reserves (Hewitt et al. 2009b, Lohrer et al. 2010). Twelve replicate cores had previously been collected at Reef and Shoal Bay (October 2009, AC Central Waitemata ecological monitoring programme). These sites were heterogeneous shelly-sand sites and Reef was partially covered by seagrass.

Using PRIMER v.6 software, we developed species accumulation curves using all available samples from each site. Species accumulation curves use randomization techniques to show how the number of identified taxa increases with the number of samples collected irrespective of how the individual samples are ordered. Because the SUMactual value (and ultimately the NIWACOOBII) relies on the number of taxa in seven functional trait groups, we examined the rates at which taxa accumulated in those trait groups at each site (Top, Erect, Surface-to-Surface, Sedentary, Suspension Feeder, Medium, Worm-like). The species accumulation procedure provided an average number of taxa expected per replicate (+ standard deviation) for all sample sizes up to the maximum number of cores collected at a particular site. For each site, we created a table of average values and standard deviations for 3, 6, 9, 10 and 12 replicates. Each standard deviation was multiplied by 2 and added to its corresponding average (in a normal distribution, this approximates the 95^{th} percentile value). For each site and number of replicates, this approximate 95^{th} percentile value was then multiplied by 1.05 (to make it 5% larger), hopefully ensuring that this value would be larger than all actual observed values without being unrealistically large (e.g., double or triple the magnitude).

For each level of replication, the highest value across the four sites for “Top” was selected as $Ntaxa_{TopMAX}$. The highest value across the four sites for “Erect” was selected as $Ntaxa_{ErectMAX}$, and so on, until all seven $Ntaxa_{groupMAX}$ values for a given level of replication (3, 6, 9, 10, 12) were available. The sum of the seven $Ntaxa_{groupMAX}$ values for a given level of replication was accepted as SUMmax for that level of replication.

2.3 Data for NIWACOOBII testing

Macrofauna, mud content and heavy metal concentration data were gathered together from 34 North Island sites. Data from n>9 macrofauna replicates were available at all 34 sites. In contrast, similar to the initially used Regional Discharges Project sites, mud and heavy metal data were generally available as an average of one to three bulked sediment cores per site.

The sites used in this analysis were originally sampled for various reasons:

[2] Note, collecting the same number of replicates per site (e.g., 12) when the sites have widely differing sampling areas (e.g., 100 m x 100 m versus 30 m x 60 m) could also influence taxonomic richness and therefore the NIWACOOBII calculations. However, exploring the bias of site area is beyond the scope of the current investigation.

- The Waiwera samples were associated with the AC estuarine monitoring programme, which usually involves the collection of n=6 macrofaunal replicates from multiple sites per estuary at multiple estuaries. We collected and analysed an additional 4 replicates per site at Waiwera in October 2010, resulting in n=10 replicates at 7 individual sites.

- Sites in the Wairoa embayment were sampled in association with an AC Tier II mapping programme, with n=12 replicates per site collected at seven intertidal sites in October 2010.

- Data from the AC Upper Waitemata Harbour ecological monitoring programme (12 sites, n=12 macrofaunal replicates per site, sampled in Nov 2009) were also used. These sites were from the same part of the Harbour but sometimes from different tidal creeks or parts of tidal creeks. Sediments were muddy (>20%) at ten of the twelve sites, and heavy metal concentrations were high (Cu 15-20, Pb 25-30, Zn 75-100 mg/kg dry wt) at all twelve sites.

- Data from an additional 8 sites had previously been collected as part of NIWA's FRST-funded Eco-Diagnostics project. The sites were Coxes Bay, Hobson Bay, Tamaki Estuary (at Panmure), Taylors Bay (also known as Hillsborough), Huia, Whitford, Pollen Island, and Ahuriri Estuary. At each of the sites, n=9 macrofaunal replicates were collected (April/May of 2006-2008). Unlike Wairoa and Waiwera, the EcoDiagnostics sites were separated by many kilometers (the Ahuriri site was in Napier) and in different water bodies (Manukau and Waitemata Harbours, Tamaki Estuary and Strait).

The heavy metals included in our analysis were total recoverable Cu, Pb and Zn concentrations in the <500 µm sediment fraction (RJ Hill Laboratories Ltd). From the individual metal concentrations, we calculated a combined metal index called CCU (Cumulative Criterion Unit, Clements et al. 2000). We also converted individual metal concentrations to PCA1.500 scores[3]. PCA1.500 scores were first introduced during the development of the Benthic Health Model (Anderson et al. 2002, 2005) and were used subsequently as a metric of heavy metal contamination (Hewitt et al. 2005, 2009a, Thrush et al. 2008) and during NIWACOOBII development (van Houte-Howes and Lohrer 2010).

All macrofaunal samples were identified in the NIWA Hamilton bio-laboratory with taxonomy resolved to genus and species whenever possible. Many groups are not well described in New Zealand and resolution to family (or even phylum) was reasonably common. However, in such cases, care was always taken to differentiate forms that appeared to be distinct morpho-species even if the exact taxonomic names were not known (e.g., "Lepidonotinae" was not the same as "*Lepidonotus polychromus*", which was not the same as "Polynoid B" (which is now known as *Paralepidonotus ampulliferus*)).

[3] $PCA1.500 = 0.615*(\log_e [Cu] - 2.472) + 0.528*(\log_e [Zn] - 4.418) = 0.586*(\log_e[Pb] - 2.925)$

2.4 NIWACOOBII calculations

To calculate the NIWACOOBII index, the list of taxa found in a particular set of samples (i.e., the 10 replicates from Wairoa Site 1) was matched to the functional traits database. When a species was found to be present at a site, we considered that its "functional trait" was present at the site, even if the species may have performed the function less than 100% of the time. For example, many deposit feeders will switch to suspension feeding when conditions are suitable[4]. The presence of such a species at a site adds a "1" to the SUMactual total by contributing a "1" to the $Ntaxa_{Sus}$ quantity (the number of taxa in the "suspension feeders" trait group; see section 1.3). Note that although the species is also counted in the deposit feeder group, the richness of deposit feeders ($Ntaxa_{Dep}$) does not factor into the SUMactual equation (see section 1.3). Calculations like this were made for all of the taxa observed in the data set and across all of the seven NIWACOOBII trait groups. Once finished, the SUMactual total could be tallied.

The SUMactual value was then used in the NIWACOOBII formula:

$$1 - (SUMmax_n - SUMactual_n)/SUMmax_n$$

If SUMactual was derived from *n* macrofaunal replicates, a SUMmax value appropriate for a sample size of *n* was used along with it. A table of $SUMmax_n$ values for samples sizes of 3, 6, 9, 10, and 12 is presented in the Results section.

2.5 Taxonomic resolution and NIWACOOBII scores

We calculated NIWACOOBII scores at the Pollen Island and Waiheke Island sites using lists of taxa identified to the lowest practicable level (mainly genus and species). We then took the same list and assumed that we could not distinguish organisms below the family level. All organisms from the same family were considered as one indistinguishable "species". Occasionally this resulted in the need to apply fuzzy probabilities to determine functional traits. We then re-calculated the NIWACOOBII scores and compared them to the NIWACOOBII scores from the fully resolved taxonomic list.

[4] Many spionid polychaetes and also some ophiuroids (e.g., *Amphiura* spp.) will switch from deposit feeding to suspension feeding during immersion periods with peak tidal flow, as particle capture rates make the switch beneficial from an energetics standpoint.

3 Results

3.1 The functional traits database

To reiterate, the NIWACOOBII is generated from 7 functional trait groups. The index tally goes up any time an organism fits into one or more of the 7 functional trait groups and is present at a site. Organisms that do not fit into any of the 7 groups obviously make no contribution to the NIWACOOBII index value, although these types of organisms appear to be very rare. Only ten of the 481 taxa listed in the functional traits database (2%) did not fit into at least one of the seven NIWACOOBII groups. These non-contributing taxa included a rare type of parasitic gastropod, a nudibranch, a marine arachnid, small commensal crabs (pea crabs, *Pinnotheres*), a cushion star (*Patiriella*, which is rarely collected in small diameter intertidal sediment cores), and three predominantly subtidal species (the decorator crabs, *Notomitrax* and *Leptomitrax*, and the bivalve, *Offadesma angasi*).

On the other hand, 98% of the taxa listed in the functional traits database belonged to at least one of the seven NIWACOOBII categories (Table 2). This means that nearly every taxon identified at a site will contribute at least a “1” to the NIWACOOBII tally[5]. Furthermore, 120 taxa (25% of the total list) belonged to four or five NIWACOOBII categories and will thus contribute “4” or “5” to the overall NIWACOOBII tally when present at a site. The types of taxa that will make the largest contributions to NIWACOOBII scores (“5” or “6”) are moderately sized, worm-shaped animals living in fixed tube structures extending through the upper 0-2 cm of the sediment column. Examples include Phoronida and Annelida (polychaetes from the families Terebellidae, Trichobranchidae, Pectinariidae, Ampharaetidae, Maldanidae, Sabellidae, and Oweniidae). It is conceivable that these taxa may be the most sensitive to mud and metals, given the way in which the seven categories were selected for inclusion in the NIWACOOBII, although this has not yet been tested or confirmed. These types of worms are not those normally utilised in eco-toxicological studies.

Approximately 45% of the taxa listed in the functional traits database belonged to one or two NIWACOOBII categories. These species will influence the NIWACOOBII tally, but will not have a disproportionately large effect like the tube dwelling worms listed above. Generally speaking, the types of species that belonged to just one or two categories were organisms such as gastropods as these are not sedentary, not suspension feeders, not worm-shaped and they do not create erect topographic features (tube structures) on the sediment surface.

[5] It also suggests that the NIWACOOBII will be positively correlated with taxonomic richness (i.e., higher taxonomic richness, higher NIWACOOBII score).

Table 1:

The number of taxa listed (upper row) and percentage of the total taxa listed (bottom row) that were recorded as belonging to 0, 1, 2, 3 etc. of the 7 NIWACOOBII functional trait groups. The seven groups were Top, Erect, Surface-to-Surface, Sedentary, Suspension feeding, Medium-sized and Worm-shaped (see text for definition).

	0 groups	1 group	2 groups	3 groups	4 groups	5 groups	6 groups	All 7 groups
Number of taxa (of 481)	10	76	139	128	96	24	8	0
Percent of total (of 481)	2%	16%	29%	27%	20%	5%	<2%	0%

3.2 $SUMmax_n$

The manner in which the number of taxa observed in each of the 7 functional trait groups increased with increasing sample size is shown for four sites (Figs. 1-4). At Waiheke Island (Fig. 1), the number of taxa in the 7 trait groups did not appear to saturate (reach an asymptote) even after 48 cores were sampled. The same was generally true for Pollen Island (Fig. 2), although to a lesser degree, particularly for the "Erect" and "Medium" trait groups. The Pollen and Waiheke Island sites had similar numbers of species (55 and 52, respectively) with identical amounts of replication, but Waiheke was apparently more heterogeneous with greater among-core dissimilarity (Hewitt et al. 2009b). This would tend to increase the number of samples required to saturate the sampling of the site's species richness.

At Reef and Shoal Bay (Oct 2009), there were more taxa in 12 replicates (67 and 63, respectively) than there were in 48 replicates at Pollen and Waiheke. There was no evidence that the number of taxa in any of the 7 functional trait groups had reached an asymptote after 12 replicates (Figs. 3-4). It is important to note that given the diversity of coastal marine benthic communities this is not unusual (Thrush et al. 1988, Gray 2002).

The estimated maximum values for each of the seven functional trait groups are shown in Table 3. Results for five different levels of replication (n=3, 6, 9, 10, 12) are also given. $SUMmax_n$ values for each level of replication, n, were calculated by summing across the 7 different functional trait groups. These individual trait group maximum values ($Ntaxa_{groupMAXn}$) and the $SUMmax_n$ values are also presented graphically in Fig. 5.

Figure 1:

The predicted (average ± 1 standard deviation) numbers of taxa (richness) accumulating in each of the 7 functional trait groups at the Waiheke Island site. Note the scale change for each panel.

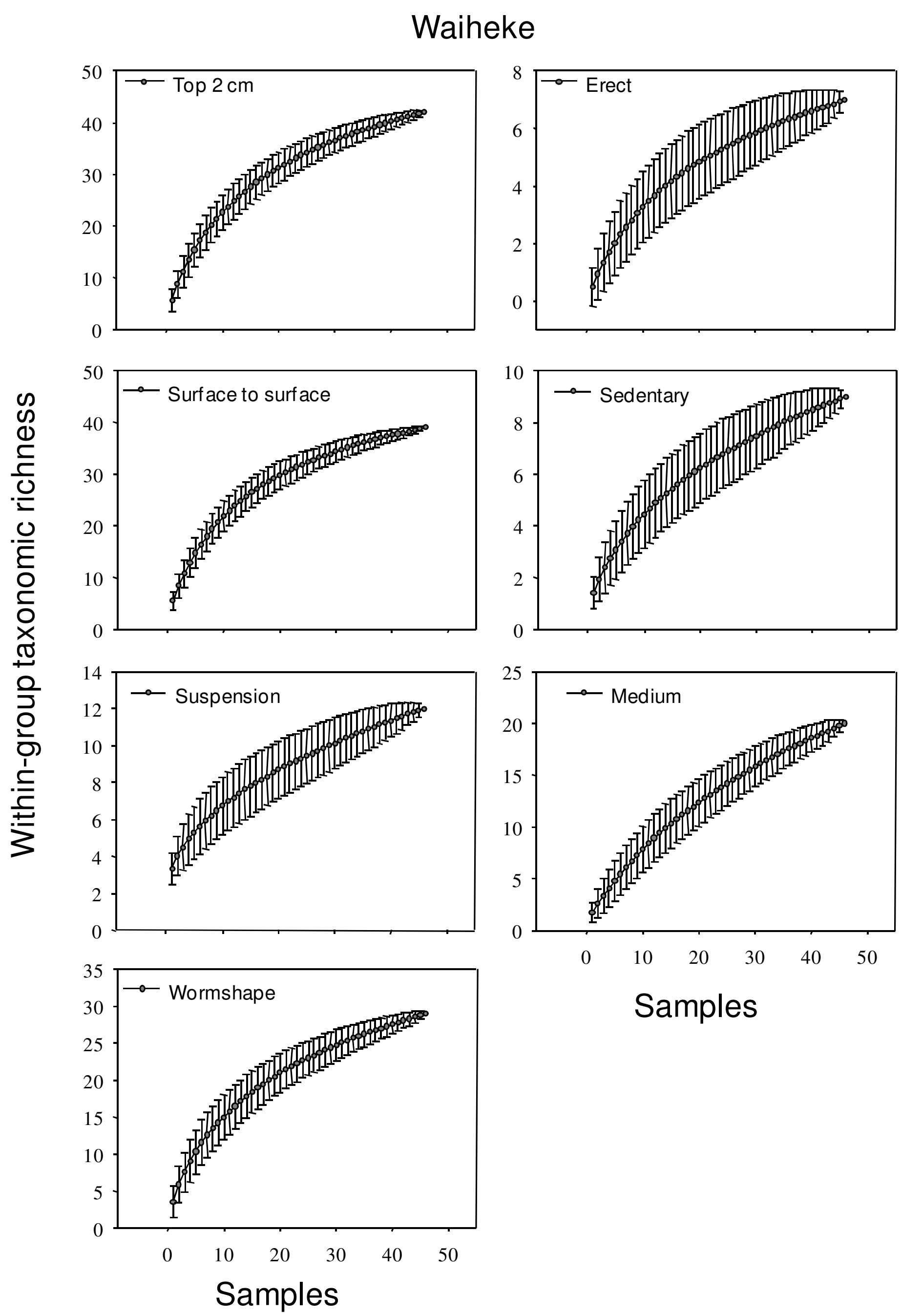


Figure 2:
The predicted (average ± 1 standard deviation) numbers of taxa (richness) accumulating in each of the 7 functional trait groups at the Pollen Island site. Note the scale change for each panel.

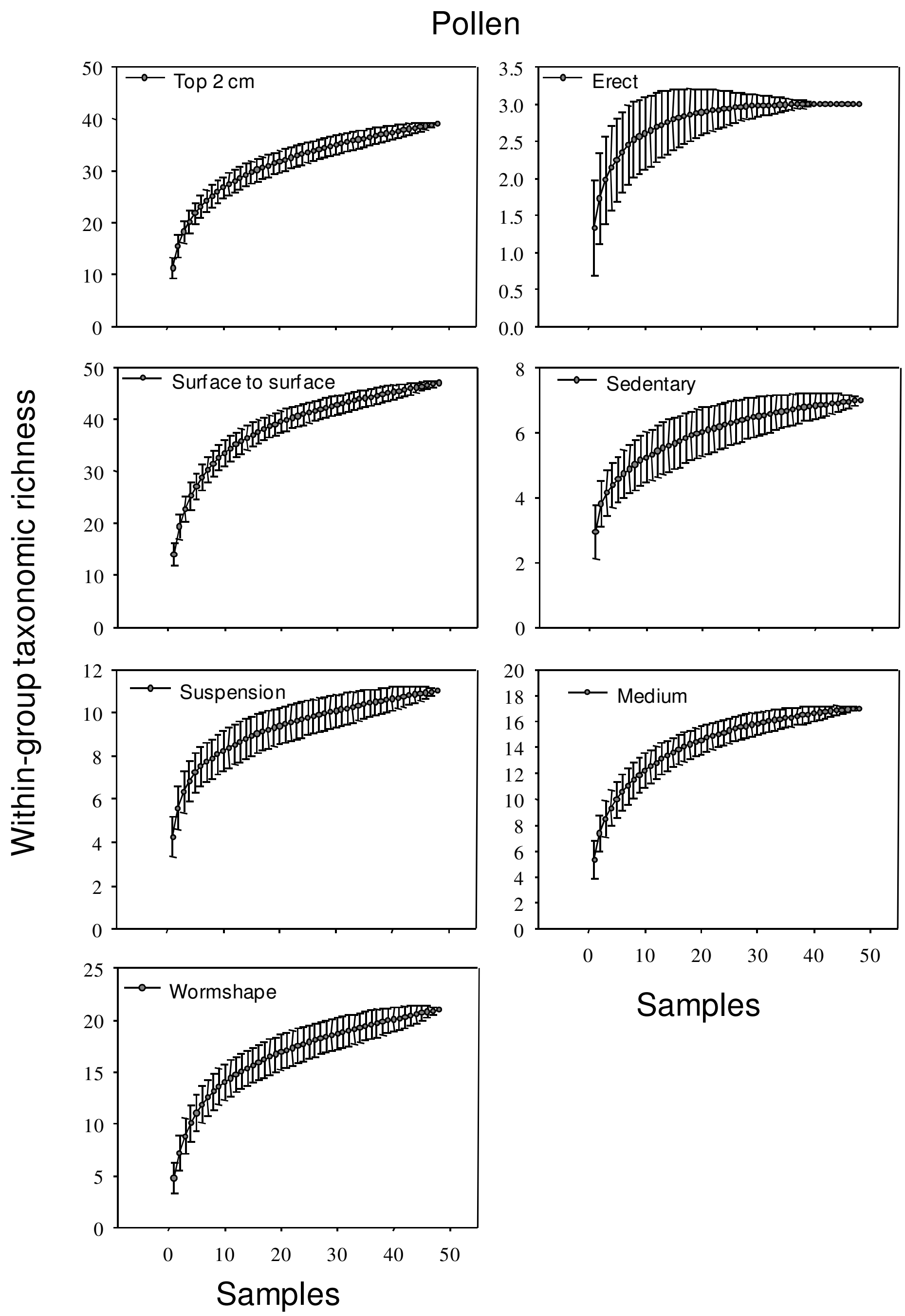


Figure 3:
The predicted (average ± 1 standard deviation) numbers of taxa (richness) accumulating in each of the 7 functional trait groups at the Reef site (Central Waitemata Harbour, October 2009). Note the scale change for each panel.

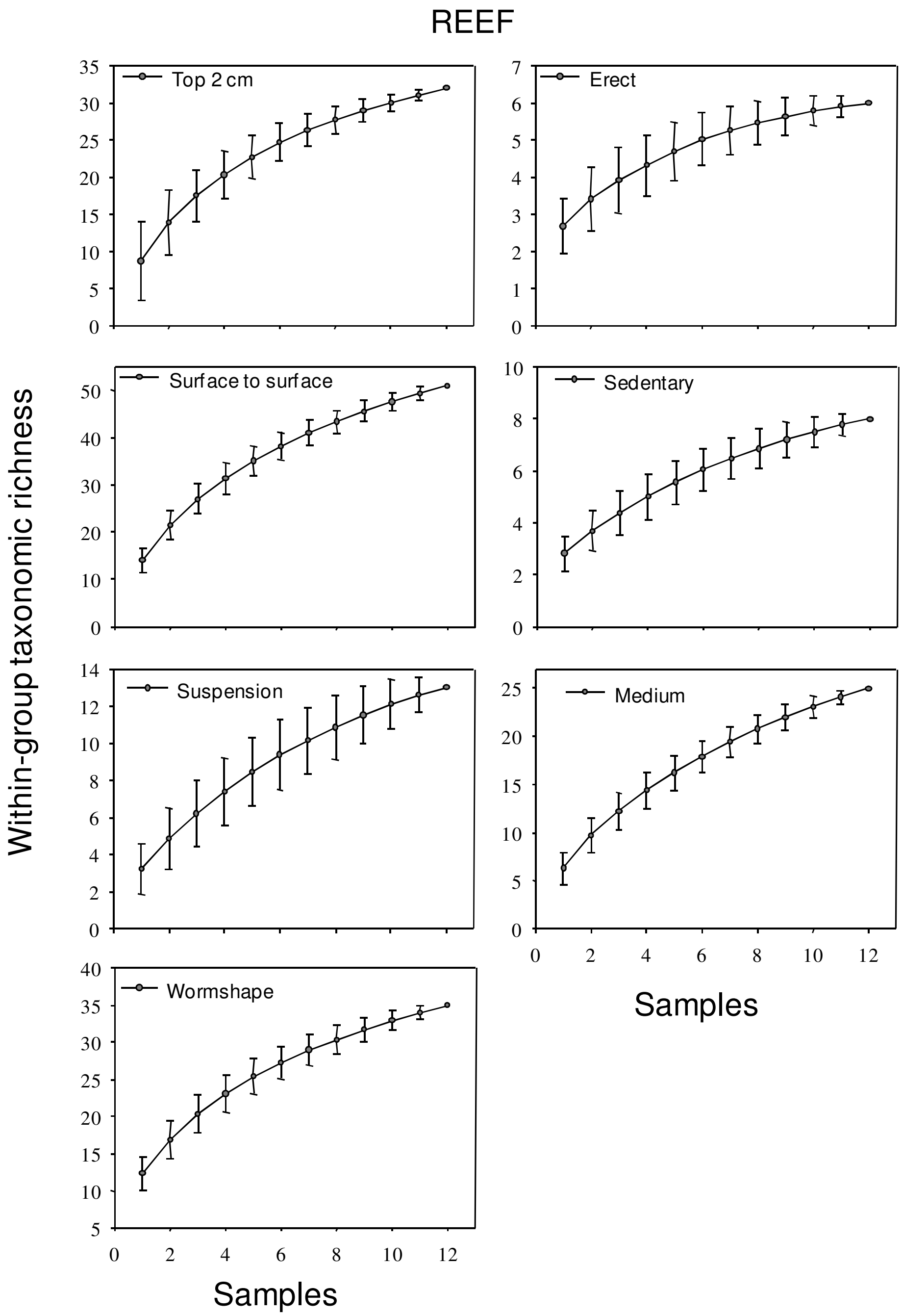


Figure 4:

The predicted (average ± 1 standard deviation) numbers of taxa (richness) accumulating in each of the 7 functional trait groups at the Shoal Bay site (Central Waitemata Harbour, October 2009). Note the scale change for each panel.

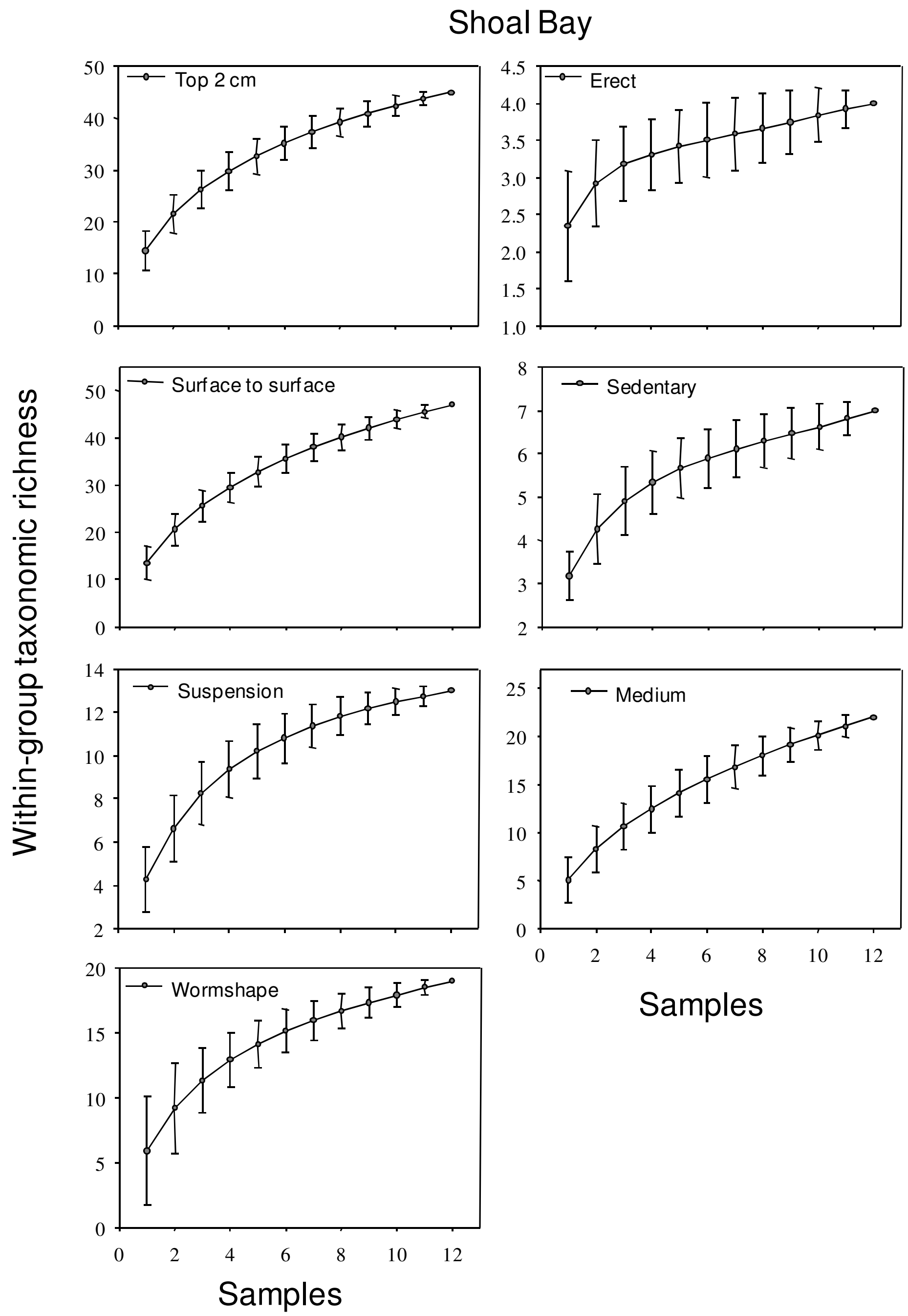


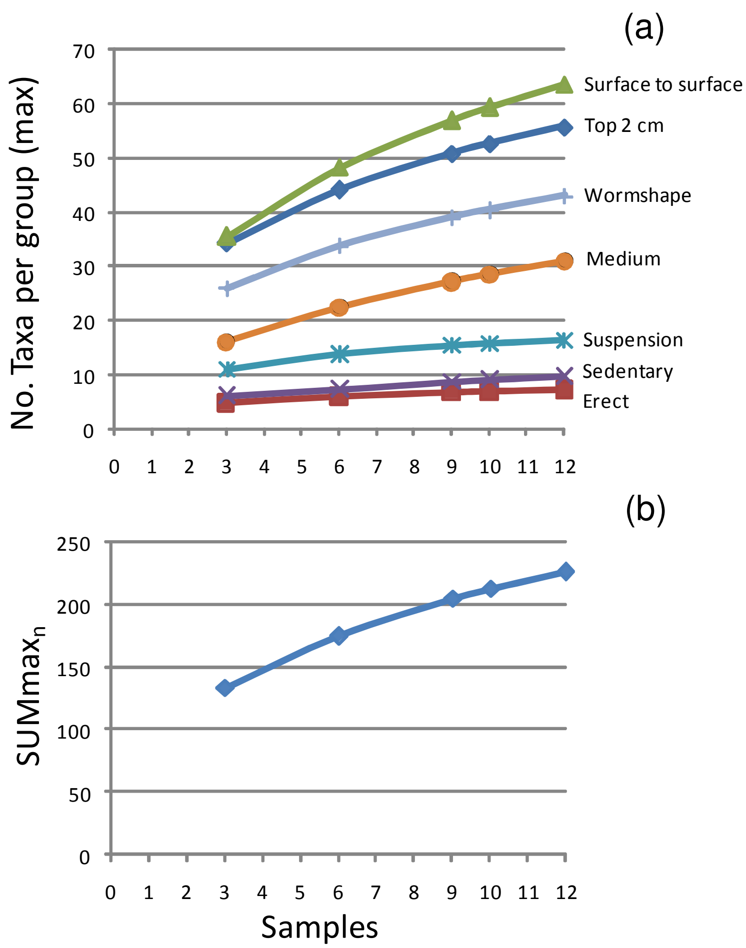
Table 2:

Estimated theoretical maximum number of taxa for each of the seven functional trait groups, along with the overall $SUMmax_n$ values, for differing levels of replication.

No. of reps	Top	Erect	SS	Sed	Sus	Medium	Worm	**$SUMmax_n$**
3	34.24	4.81	35.49	6.15	10.90	16.15	25.84	**133.56**
6	44.03	6.02	48.05	7.33	13.81	22.44	33.66	**175.35**
9	50.70	6.81	56.78	8.71	15.42	27.25	38.91	**204.59**
10	52.48	7.01	59.17	9.09	15.81	28.59	40.36	**212.51**
12	55.63	7.33	63.38	9.76	16.44	30.97	42.89	**226.39**

Figure 5:

The number of replicates collected at a site along the x-axis plotted versus (a) the theoretical maximum number of taxa per group for each of the 7 functional trait groups and (b) the $SUMmax_n$ parameter.



3.3 NIWACOOBII calculations

Using the updated functional traits database and $SUMmax_n$ values, NIWACOOBII calculations were made at 34 sampling sites. The NIWACOOBII scores are presented for each site, along with the sediment mud and heavy metal data (Table 4). Given that the NIWACOOBII is an index based on taxonomic richness in seven functional trait groups, we also present total taxonomic richness at each site for comparative purposes.

The data set we compiled had some unique features. Most notably, sites HIW and HIN in the Upper Waitemata Harbour had low mud content (<10%) but very high metal concentrations. Overall, as mud content increased beyond 20%, metal contaminant concentrations appeared to plateau. This is evident when sediment mud content is plotted versus PCA1.500 (an index of heavy metal contamination based on summed Cu, Pb and Zn concentrations) (Fig. 6).

Figure 6:
Relationship between sediment mud content (mud %) and sediment heavy metals (PCA1.500). PCA1.500 is an index of heavy metal contamination based on a combination of Cu, Pb and Zn concentrations. High PCA1.500 scores indicate high levels of metal contamination. Data from 34 sites are included in this analysis.

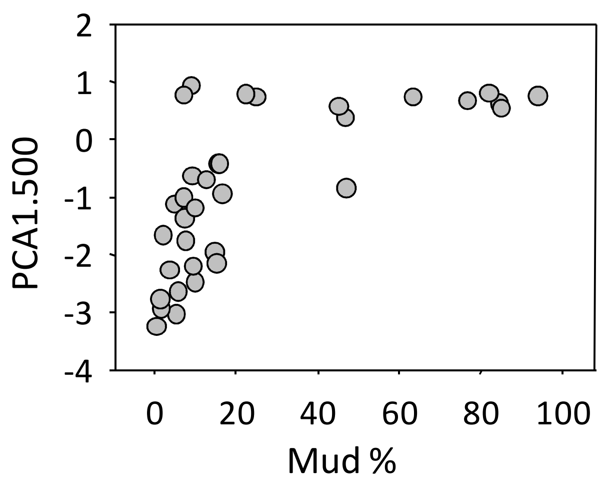


Table 4:

A listing of sediment characteristics including percent mud content and concentrations of copper, lead and zinc (mg/kg dry wt). Two combined metal indices are also listed (CCU and PCA1.500). The last 5 columns on the right hand side refer to the number of macrofaunal samples collected per site (n), the corresponding $SUMmax_n$ values used (taken from Table 3), the calculated $SUMactual_n$ totals (based on n replicates), the resultant NIWACOOBII scores, plus overall taxonomic richness values. Note that the NIWACOOBII index, with the formula $1 - (SUMmax_n - SUMactual_n)/SUMmax_n$, is comparable across sites with differing levels of replication (n) and varies between 0.0 and 1.0.

Site	Mud %	Cu	Pb	Zn	CCU	PCA 1.500	n	SUM max_n	SUM $actual_n$	NIWACOOBII	Richness
Waiwera1	47.2	8.0	9.4	45.0	0.86	-0.84	10	212.5	64	0.30	20
Waiwera2	14.9	4.0	4.1	31.0	0.48	-1.95	10	212.5	35	0.16	12
Waiwera3	0.5	1.5	1.8	21.0	0.26	-3.24	10	212.5	56	0.26	20
Waiwera5	9.4	4.0	2.9	28.0	0.42	-2.21	10	212.5	119	0.56	41
Waiwera6	1.6	2.0	2.0	24.0	0.31	-2.93	10	212.5	91	0.43	35
Waiwera8	1.4	2.0	2.7	23.0	0.32	-2.78	10	212.5	105	0.49	39
Waiwera10	15.3	5.0	2.8	25.0	0.42	-2.15	10	212.5	83	0.39	30
Wairoa1	7.6	3.0	6.0	41.0	0.59	-1.75	12	226.4	72	0.32	31
Wairoa2	10.0	1.5	5.2	27.0	0.41	-2.49	12	226.4	115	0.51	43
Wairoa3	5.7	1.5	4.3	25.0	0.37	-2.64	12	226.4	106	0.47	43
Wairoa4	2.1	2.0	8.5	53.0	0.73	-1.66	12	226.4	118	0.52	42
Wairoa5	3.8	1.5	5.5	39.0	0.52	-2.26	12	226.4	108	0.48	40
Wairoa6	16.7	5.0	11.3	53.0	0.89	-0.93	12	226.4	113	0.50	38
Wairoa7	10.0	4.0	10.2	48.0	0.79	-1.18	12	226.4	95	0.42	32
Brig	84.8	18.6	25.7	87.7	1.97	0.62	12	226.4	29	0.13	9
RNG	94.2	20.1	26.7	98.0	2.12	0.75	12	226.4	81	0.36	26
Hell	46.7	14.1	26.0	75.0	1.75	0.38	12	226.4	43	0.19	14
Luc	45.3	16.1	26.0	93.3	1.95	0.57	12	226.4	89	0.39	30
HIN	9.1	21.2	32.0	105.0	2.36	0.92	12	226.4	85	0.38	33
MainU	85.1	17.1	24.7	87.7	1.89	0.54	12	226.4	44	0.19	15
OHBV	76.8	16.5	29.0	96.0	2.07	0.66	12	226.4	87	0.38	29
MainC	25.0	17.9	29.7	97.3	2.14	0.74	12	226.4	74	0.33	24
MainO	22.5	19.7	28.3	99.3	2.17	0.78	12	226.4	76	0.34	28
HellU	82.1	18.0	30.7	104.7	2.23	0.80	12	226.4	49	0.22	16
LucU	63.5	18.4	28.0	99.7	2.13	0.73	12	226.4	97	0.43	32
HIW	7.2	18.9	28.7	100.0	2.16	0.77	12	226.4	77	0.34	36
Coxes Bay	4.9	3.7	10.6	56.0	0.86	-1.13	9	204.6	81	0.40	30
Hobson Bay	9.1	4.8	19.0	55.3	1.13	-0.63	9	204.6	54	0.26	25
Huia	7.1	8.9	6.0	47.3	0.81	-1.01	9	204.6	83	0.41	31

Site	Mud %	Cu	Pb	Zn	CCU	PCA 1.500	No. reps	SUM max_n	SUM $actual_n$	NIWACOOBII	Richness
Pollen "B"	7.4	4.1	9.9	34.0	0.67	-1.37	9	204.6	71	0.35	29
Taylors Bay	15.7	6.9	15.0	70.7	1.20	-0.42	9	204.6	86	0.35	30
Whitford	5.4	1.6	3.0	16.0	0.26	-3.04	9	204.6	71	0.52	42
Ahuriri	16.0	7.8	12.7	71.7	1.17	-0.43	9	204.6	107	0.26	18
Pollen	5.1	8.1	15.6	42.8	1.10	-0.56	48	226.4	53	0.64	55
Waiheke	8.6	8.6	7.9	17.7	0.57	-1.40	48	226.4	145	0.70	52
Reef Oct09	10.1	n.d.	n.d.	n.d.	n.d.	n.d.	12	226.4	158	0.83	64
ShB Oct09	9.8	n.d.	n.d.	n.d.	n.d.	n.d.	12	226.4	188	0.73	60

Across the 34 sites, NIWACOOBII scores declined significantly with increasing amounts of mud and heavy metals (Fig. 7, Table 5). On average, NIWACOOBII scores declined by about 60% (from 0.41 to 0.25) across the entire range of mud content values (0-94% mud). However, the amount of variation in NIWACOOBII scores attributable to mud content was rather low (r^2 = 0.232).

Decreases in NIWACOOBII scores with increasing heavy metal concentrations were observed for Cu, Pb and Zn individually and for the metal gradient indices CCU and PCA1.500 (Fig. 7, Table 5). However, as with mud, the percent variability explained by these predictor variables was low ($0.162< r^2 < 0.218$). We doubt that the low r^2 reflects confounding by an unmeasured underlying gradient, such as salinity. Rather, the number of sites in the data set was relatively small and the sites were not specifically selected for the breadths of their mud and metal gradients. All of the sites with greater than 20% mud (as well as HIN and HIW with <10% mud) had relatively high heavy metal concentrations (Fig.6), which likely affected richness and thereby compressed the amount of variability that could be explained by either variable by itself. It is also important to note that richness and related indices (e.g., the NIWACOOBII) cannot account for species turnover, for example the replacement of sensitive taxa by less sensitive taxa as a result of increasing contaminants, even though this is known to occur. Community-based analyses are better for detecting such shifts, which is why the BHM is able to explain a much higher proportion of variability than richness or the NIWACOOBII.

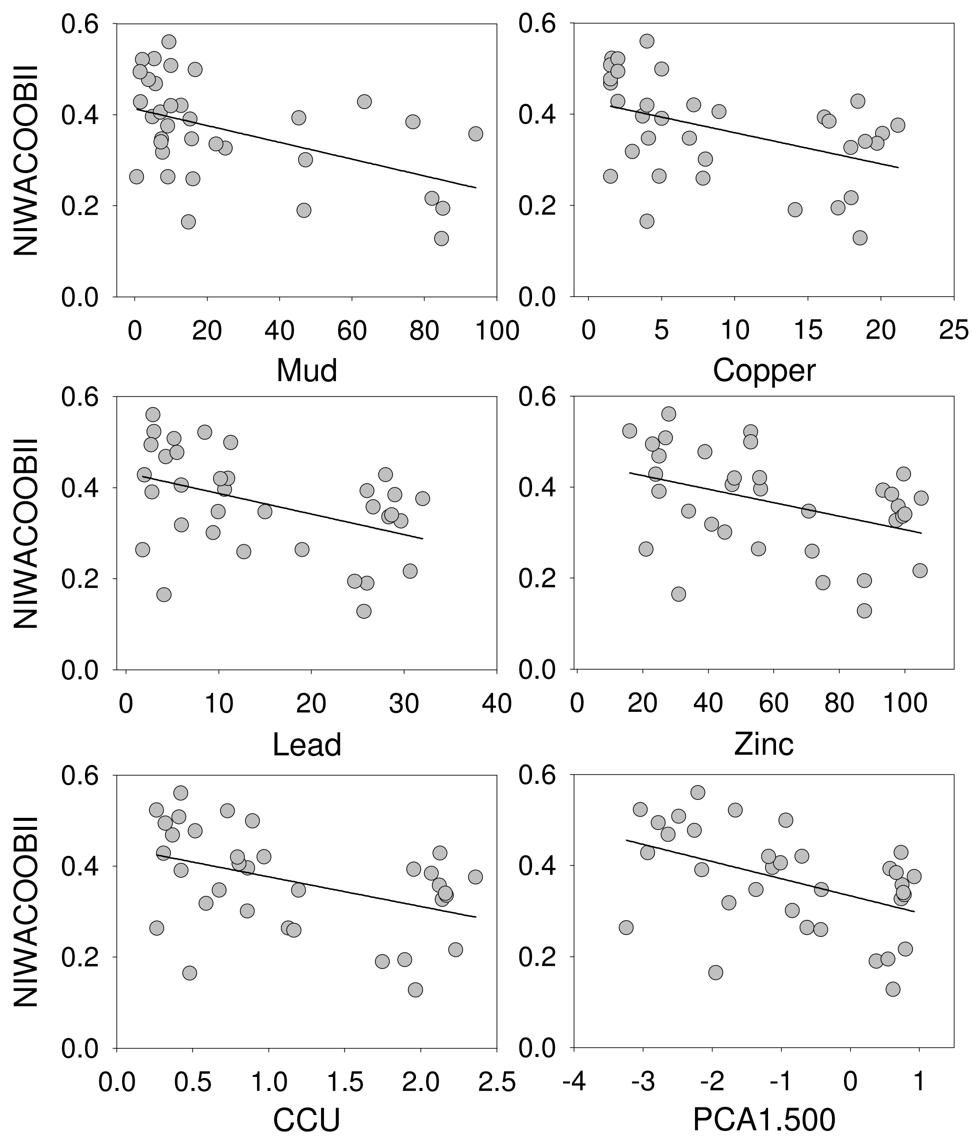
Table 4:

Results of data analysis including the intercept and slope from linear least-squares regression fits of predictor versus response; r^2 values and significance levels (* p < 0.05, ** p < 0.001, ** p < 0.0001) are also presented for each model.

Response Variable	Intercept	Slope	Predictor Variable	r^2	P
NIWACOOBII	0.413	-0.002	Mud	0.232	**
	0.428	-0.01	Copper	0.194	**
	0.432	-0.005	Lead	0.19	**
	0.455	-0.001	Zinc	0.162	*
	0.441	-0.0065	CCU	0.187	*
	0.333	-0.04	PCA1.500	0.218	**
Richness	34.02	-0.185	Mud	0.323	***
	34.55	-0.583	Copper	0.194	**
	34.56	-0.369	Lead	0.166	*
	36.6	-0.122	Zinc	0.151	*
	35.5	-5.33	CCU	0.173	*
	26.5	-3.2	PCA1.500	0.216	**

Figure 7:

The response of the NIWACOOBII functional traits index (which ranges from 0.0 to 1.0) to sediment mud content and sediment heavy metal concentration. Higher NIWACOOBII scores indicate healthier sites. All relationships were significant at $p < 0.05$; r^2 values were less than 0.25.



NIWACOOBII scores were expected to be correlated with overall taxonomic richness, and they were (r^2 = 0.911, Fig. 8). As a comparison to the performance of the NIWACOOBII, we plotted overall taxonomic richness versus the same set of predictor variables (mud%, Cu, Pb, Zn, CCU, PCA1.500; Fig. 9).

Overall taxonomic richness had a slightly stronger relationship with sediment mud percentage (r^2 = 0.323, relative to r^2 = 0.232 for NIWACOOBII), whilst the NIWACOOBII tended to have slightly stronger relationships with the metals gradients (see r^2 values for Pb, Zn, CCU, PCA1.500 in Table 5). None of the relationships was particularly strong, with <25% of the variation explained in all cases except one.

Figure 8:
The positive relationship between NIWACOOBII score and overall taxonomic richness.

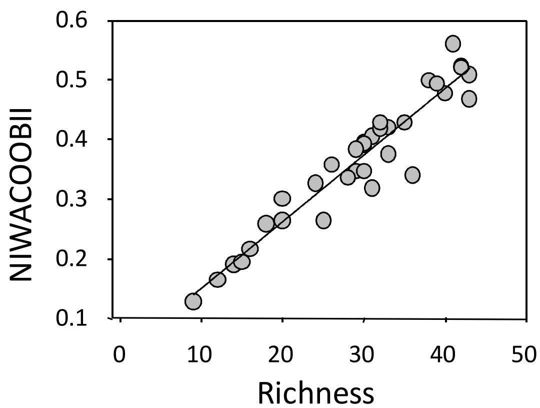
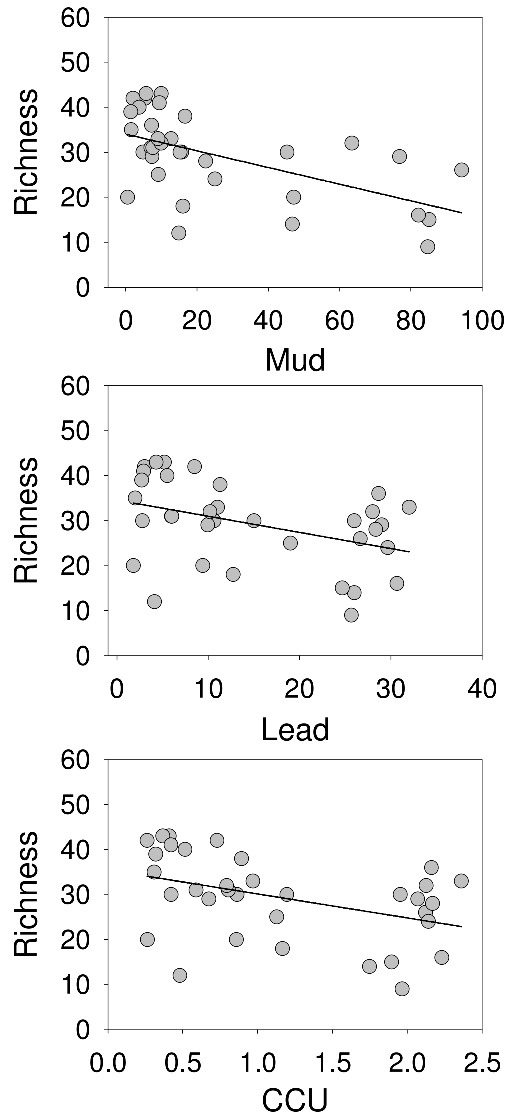


Figure 9:

The response of overall taxonomic richness to sediment mud content and sediment heavy metal concentration. All relationships were significant at $p < 0.05$; r^2 values were less than 0.25, with the exception of the upper left panel (mud, $r^2 = 0.323$).



Mud content and heavy metal concentrations are known to be correlated, and it is often difficult to separate out their independent effects. The NIWACOOBII was originally developed with this in mind, using a mud gradient from an area largely unaffected by metal contaminants (Mahurangi Harbour) and a contaminant gradient from sites selected to balance a range of contaminants across a mud gradient (Regional Discharges Project).

To examine this further, we selected, from the 34 new sites, a subset of 25 sites that had a limited range of sediment mud content values of 0-17% (cf. the 0-94% range across all 34 sites). The corresponding range of heavy metal concentrations at the 25 sites was larger (e.g., PCA1.500 scores from -3.24 to +0.92), but still restricted as the most contaminated sites had contaminant concentrations well below current guideline values. For these 25 sites, there was no significant relationship between any of the metal variables and the NIWACOOBII.

We also examined a subset of 12 sites, all from the Upper Waitemata Harbour, that had a broad range in sediment mud content (7 – 94%) and a narrow, although relatively high, range of metal contaminant values (PCA1.500 from +0.38 to +0.92). At these sites, we regressed sediment mud content versus NIWACOOBII scores. The relationship between NIWACOOBII scores and mud was not significant, although the low number of sites (n = 12) and the higher level of contamination at these sites both work against a significant relationship being observed.

3.4 Taxonomic resolution and NIWACOOBII scores

When species were identified to the lowest possible level (genus and species, rather than family only), this resulted in the enumeration of 17 additional taxa at Pollen Island and 12 additional taxa at Waiheke Island. The greater level of taxonomic differentiation had marked effects on SUMactual values as well as on taxonomic richness within most of the individual functional trait groups (Table 6). The implication is that poor taxonomic differentiation resulting in artificially low SUMactual values will result in low NIWACOOBII scores, which will lead to incorrect conclusions about the level of degradation at sites of interest.[6]

[6] Technically, $SUMmax_n$ values could be re-adjusted downward based on family-level accumulation curves (e.g., adapted methods similar to Section 2.2), but we do not recommend this course of action.

Table 6:

The number of taxa in each of the NIWACOOBII functional trait categories is presented for two well sampled sites (Pollen, Waiheke, n = 48 each). The number of taxa in each group was tabulated after differentiating organisms to family level and then again after differentiating organisms to the lowest taxonomic level possible (genus and species).

	Genus & species Pollen	Family Pollen	Genus & species Waiheke	Family Waiheke
Top	39	27	42	32
Erect	3	2	7	4
SS	47	33	39	32
Sedentary	7	6	9	6
Suspension	11	10	12	11
Medium	17	14	20	16
Wormshape	21	13	29	19
SUMactual	**145**	**99**	**158**	**114**
Richness	55	38	52	40

4 Discussion

Indicators of ecological integrity are designed to condense complex information into a form that is suitable for presentation in a simple non-technical way. The Auckland Council possesses extensive data sets and detailed ecological information from monitoring programmes and other research projects, and it seeks ways of presenting this information clearly and concisely to facilitate State of Environment reporting and communication to the general public.

Although a number of ecological indicators have been developed overseas, these do not appear to be readily transferable to New Zealand data sets. When tested, two commonly used overseas indices failed to respond to strong gradients of mud and heavy metals across 95 sites in the Auckland area (van Houte-Howes and Lohrer 2010). Thus, indicators of ecological integrity more relevant to New Zealand's environmental issues and habitat types are specifically sought by the Auckland Council.

The NIWACOOBII is a New Zealand based index that was recently developed as a proof of concept (van Houte-Howes and Lohrer 2010). The NIWACOOBII index is calculated based on the richness of macrofaunal organisms in seven functional trait groups. The seven functional trait groups used in the calculation were not selected for their "meaningfulness" from an ecosystem functioning standpoint, but rather for their sensitivity to mud and heavy metals (which are two primary stressors influencing macrofaunal community structure in Auckland Region intertidal estuarine habitats; Hewitt et al. 2005, 2009a; Thrush et al. 2008).

Although not yet fully refined in 2010, the NIWACOOBII was better correlated with mud and heavy metal contaminant gradients than either of the two overseas indices mentioned above. The ability of the index to track these stressors was encouraging, although adoption of the NIWACOOBII as a potential State of Environment metric was not advised until further testing and verification with independent data could be performed.

Here, prior to the testing, adjustments were made to improve the NIWACOOBII calculation. The adjustments involved changes to the SUMmax parameter. The other value involved in the calculation, SUMactual (the summed taxonomic richness in the seven specified functional trait groupings), remained essentially unchanged[7].

The SUMmax value was originally conceived to represent a theoretical maximum value, namely a reference value that is likely to be achieved only at a perfectly healthy, non-polluted site. As such, the SUMmax parameter needs to be set high enough that it is not exceeded by actually observed values (i.e., SUMactual). On the other hand, the SUMmax value cannot be set unrealistically high, as this will simply compress all the NIWACOOBII scores to the point where they are indistinguishable from one another (a clump of compressed NIWACOOBII scores will show no change in slope when plotted along contaminant gradients; cf Fig. 7). Therefore, a logical and non-arbitrary technique for setting the SUMmax parameter was needed so that the NIWACOOBII scores would be both meaningful and useful. The technique utilized involved species

[7] This means that NIWACOOBII scores calculated prior to 2011 can be recalculated and compared to new NIWACOOBII scores if so desired.

accumulation curves and, specifically, rates of species accumulation in the seven identified functional trait categories.

The use of species accumulation techniques to set SUMmax also allowed us to quantify and eliminate the influence of sample size on NIWACOOBII scores. A table of SUMmax values was given for replication levels of 3, 6, 9, 10, and 12 cores per site (Table 3). Therefore, assuming that the appropriate $SUMmax_n$ values are used for each, NIWACOOBII scores from various sites with differing levels of replication can now be validly compared.

The NIWACOOBII is currently set up under the assumption that macrofauna will be differentiated to the lowest practicable level (genus/species). For example, the $SUMmax_n$ values of Table 3 are based on fully differentiated data. SUMactual scores developed from family-level taxonomy were 30 to 40% lower than species-level scores at the two sites we tested (Table 6). Accordingly, family-level scores would not be suitable for use with the $SUMmax_n$ values of Table 3; the use of family-level taxonomy would require a re-vamp of the entire NIWACOOBII calculation. Differentiation of macrofauna to the lowest practicable level is already done at least once a year (October) in all current AC ecological monitoring programmes, which is amenable to NIWACOOBII calculations in their current form. In theory, the level of expertise in taxonomy at one laboratory versus another could affect the NIWACOOBII to some degree. Consistency in the level of species identifications over time and among sites is the best way to ensure intercomparability of NIWACOOBII values.

4.1 NIWACOOBII performance

On an independent set of 34 sites, the NIWACOOBII showed a significant response to gradients of both mud and metals. Thus, as a proof of concept, the NIWACOOBII once again showed the ability to track real (and multiple) environmental stressors. Declines in NIWACOOBII scores with increases in mud and heavy metals are interpreted as losses of functional redundancy. Habitats with high functional redundancy (i.e., many species present within each functional trait group) will tend to have higher inherent resistance and resilience in the face of environmental changes, as the higher numbers of species per functional group provide “insurance” for stochastic or stress-induced losses of particular species. Higher numbers of species per functional group probably equates to a greater range of activity types within functions as well. Therefore, the NIWACOOBII analysis is meaningful with regards to maintaining ecosystem multi-functionality.

NIWACOOBII scores were correlated with overall taxonomic richness at the 34 test sites, thus the trends in the two metrics along the gradients of mud and metals were generally similar: both declined significantly with increasing mud and metals, and neither showed marked superiority over the other in terms of percent variability explained (r^2) by mud or heavy metals. However, overall taxonomic richness was more highly correlated to mud than to heavy metals, whereas NIWACOOBII was equally well related to both.

Overall taxonomic richness is a very easy metric to obtain, but its meaning for ecosystem health has not been demonstrated. Analyses based on functional traits or indices (e.g., the NIWACOOBII), although slightly more difficult to calculate at this

stage, provide more information about the meanings of species loss—the types of species that are disappearing—than overall taxonomic richness. With information on functional traits, organisms can be categorized according to the roles that they perform rather than simply by their degree of taxonomic relatedness, and analyses can be geared towards understanding the implications of shifts in functional types. Nevertheless, the relative benefits of information gain through the use of functional traits indices, and the lack of bias towards one particular environmental stressor (mud), has to be weighed against the somewhat more arduous task of calculating them.

The improvements made to the functional traits database in preparation for this project have expanded its utility, and the NIWACOOBII is just one potential application of the database. For example, the functional traits database could also contribute to “value mapping” exercises. Different sites could be ranked according to the contributions that they make to the provisioning of ecosystem goods and services based on the richness and abundance of organisms in specific functional trait groups at those sites. The ecosystem service of water filtration/purification would be correlated with the “suspension-feeder” function; the ecosystem service of nutrient recycling would be correlated with bioturbation intensity, quantified as a combination of “mobile” + “large” + all of the “direction of particle movement” category except “surface-to-surface”; the ecosystem service of carbon sequestration would be correlated with “calcified” group; etc.

Traits based indicators such as the NIWACOOBII will also add value to the Benthic Health Model (Anderson et al. 2002, 2006, Hewitt & Ellis 2010). The Benthic Health Model (BHM) is based on significant shifts in macrobenthic community structure along pollution gradients. As it is a multivariate model, the BHM is much more sensitive to gradients of mud and metals and has higher explanatory power than the NIWACOOBII; the BHM is currently the best available tool for detecting shifts along pollution gradients. However, unlike the NIWACOOBII functional traits approach, the BHM is highly dependent on species pools and cannot yet be applied to species pools outside of the Auckland Region. The NIWACOOBII also has advantages over the BHM in terms of providing information about what the shifts in macrobenthic community structure mean in terms of ecological resistance and resilience and functionality. Thus, when used in combination, the NIWACOOBII and BHM will be highly complementary.

The BHM has been used to categorize sites on a scale from one to five depending on their degree of contamination or muddiness. Information from the functional traits database could be used to assess the types and numbers of functions that are being lost with each successive change in rank, from one to two, two to three, three to four and four to five. This type of analysis could be used to direct restoration efforts. It may turn out that restoring a site from a rank of five to four does little to recover the species responsible for maintaining key ecosystem functions, whereas the restoration of a site from a rank of three to two might make a huge difference. Hard data on the rehabilitation potential of various field sites would underpin restoration decisions and improve the efficacy of resource management.

4.2 Conclusions and recommendations

As a result of the work described here, the NIWACOOBII calculations are now more robust and can be validly compared across sites sampled with differing numbers of

replicates. Furthermore, with independent data, the NIWACOOBII has once again been shown to respond to gradients of mud and heavy metals in intertidal soft-sediment flats. It seems to be effective in tracking the simultaneous stressors of mud and heavy metals provided that a large number of sites is examined and that the gradients are relatively broad (e.g., 0 to 94% mud).

On the other hand, the low percent variability explained in the response relationships indicated the NIWACOOBII's relative lack of sensitivity, compared to the multivariate Benthic Health Model, even though the NIWACOOBII is calculated from the 7 most sensitive of the original 29 trait groups. Traits based indices based on more "meaningful" groups, in terms of their contributions to the delivery of ecosystem goods and services, may be even less sensitive.

Although the NIWACOOBII may not be suitably sensitive to be relied upon as a standalone index for State of Environment reporting, we recommend applying it in concert with the Benthic Health Model to explore the overlaps and complementarities that the two indices may provide when used together.

5 Acknowledgements

We thank the Auckland Council for funding and providing data sets for index testing. Additional data collection and processing was funded by FRST (Contract No. C01x0504). NIWA's work on resilience and ecosystem functioning, which underpinned NIWACOOBII development, was also funded by FRST (Contract No. C01X0501, C01X0502). M. Townsend provided helpful input early on during the development of this project.

6 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. Prepared by Auckland UniServices Ltd for Auckland Regional Council. *Auckland Regional Council Technical Publication 317.*

Anderson, M.J.; Hewitt, J.; Thrush, S. (2002). The development of criteria for assessing community health of intertidal flats. Prepared by NIWA (HAM2002-048) for Auckland Regional Council*. Auckland Regional Council Technical Publication 184.*

Borja, A.; Dauer, D.M.; Díaz, R.; Llansó, R.J.; Muxika, I.; Rodríguez, G.J.; Schaffner, L. (2008) Assessing estuarine benthic quality conditions in Chesapeake Bay: A comparison of three indices. *Ecological Indicators 8*: 395–403.

Clements, W.; Carlisle, D.; Lazorchak, J.; Johnson, P. (2000). Heavy metals structure benthic communities in Colorado Mountain streams*. Ecological Applications 10(2)*: 626–638.

Gray, J.S. (2002). Species richness of marine soft sediments. *Marine Ecology Progress Series 244*: 285–297.

Hewitt, J.; Anderson, M.; Hickey, C.; Kelly, S.; Thrush, S.F. (2009a). Enhancing the Ecological Significance of Sediment Contamination Guidelines through Integration with Community Analysis. *Environmental Science & Technology 43(6)*: 2118–2123.

Hewitt, J.E.; Anderson, M.J.; Thrush, S.F. (2005). Assessing and monitoring ecological community health in marine systems. *Ecological Applications 15*: 942–953.

Hewitt, J.E.; Ellis, J. (2010). Assessment of the benthic health model. Prepared by NIWA for Auckland Regional Council. *Auckland Regional Council TR2010/034.*

Hewitt, J.E.; Thrush, S.F.; Dayton, P.K. (2008). Habitat variation, species diversity and ecological functioning in a marine system. *Journal of Experimental Marine Biology and Ecology 366*: 116–122.

Hewitt, J.E.; Thrush, S.F.; Lohrer, A.M.; Townsend, M. (2009b). A latent threat to biodiversity: consequences of small-scale heterogeneity loss. *Biodiversity and Conservation 19*: 1315–1323.

Lohrer, A.M.; Hewitt, J.E.; Hailes, S.F.; Thrush, S.F.; Ahrens, M.; Halliday, J. (2010). Contamination on sandflats and the decoupling of linked ecological functions. Austral Ecology (doi:10.1111/j.1442-9993.2010.02148.x).

Thrush, S.F.; Hewitt, J.E.; Hickey, C.W.; Kelly, S. (2008). Multiple stressor effects identified from species abundance distributions: Interactions between urban contaminants and species habitat relationships. *Journal of Experimental Marine Biology and Ecology 366*: 160–168.

Thrush, S.F.; Pridmore, R.D.; Hewitt, J.E.; Roper, D. (1988). Design of an ecological monitoring programme for the Manukau Harbour. Prepared for the Auckland Regional Water Board, *Auckland Regional Authority, Report No. 7099.*

Townsend, M. (2010) Central Waitemata Harbour Ecological Monitoring: 2000-2010. Prepared by NIWA for Auckland Regional Council. *Auckland Regional Council Technical Report 2010/060.*

Townsend, M.; Marshall, B.A.; Greenfield, B.L. (2010). First records of the Australian dog whelk, *Nassarius* (*Plicarcularia*) *burchardi* (Dunker in Philippi, 1849) (Mollusca: Gastropoda) from New Zealand. *New Zealand Journal of Marine and Freshwater Research 44(4)*: 343–348.

van Houte-Howes, K.; Lohrer, A.M. (2010). State of Environment Indicators for intertidal habitats in the Auckland Region. Prepared by NIWA for Auckland Regional Council. *Auckland Regional Council Technical Report 2010/035.*

Weisberg, S.; Ranasinghe, J.A.; Dauer, D.M.; Schaffner, L.C.; Diaz, R.J.; Frithsen, A. & J.B. (1997). An estuarine benthic index of biotic integrity (B-IBI) for Chesapeake Bay. *Estuaries and Coasts 20(1)*: 149–158.