Health of Estuarine Soft-sediment Habitats: Continued Testing and Refinement of State of the Environment Indicators

December 2012 Technical Report 2012/012

Auckland Council
Technical Report 2012/012
ISSN 2230-4525 (Print)
ISSN 2230-4533 (Online)

ISBN 978-1-927169-74-2 (Print)
ISBN 978-1-927169-75-9 (PDF)

Reviewed by	Approved for Auckland Council publication by
	
Name: Megan Carbines Position: Senior Scientist, Research, Investigations and Monitoring Unit Date: December 2012	Name: Grant Barnes Position: Manager; Research, Investigations and Monitoring Unit Date: December 2012

Recommended citation:

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

Health of estuarine soft-sediment habitats: continued testing and refinement of state of the environment indicators

Judi E Hewitt
Andrew M Lohrer
Michael Townsend

National Institute of Water and Atmospheric Research Ltd

Prepared for Auckland Council

ARC12203, May 2012

Project report Number HAM2012-067

Executive Summary

The overarching objective of the current study was to test the compatibility, comparability and complementarity of the macrofaunal-based Benthic Health Model (BHM) and Traits- Based functional Index (TBI) output and to assess the suitability of these health indicators for State of the Environment reporting. Other objectives included: comparisons of index findings to trends in contaminants from the contaminant monitoring programme; investigation of whether recalibration of BHMmetals is required due to the changing relationships between the metal contaminants; and investigation of the effect of mangrove encroachment and channel formation on the results of the BHM models.

The BHM is extremely robust and addresses changes in benthic community structure associated with gradients in mud and heavy metal contaminants. Separate models have been created for two of the key environmental contaminants in Auckland area estuaries: sediment mud content (percent silt+clay, CAPmud) and sediment heavy metal concentrations (copper, lead and zinc; CAPmetals). The TBI is a functional traits based index developed for use in AC's State of the Environment reporting. This index is based upon the richness of macrofaunal taxa in seven individual functional trait groups.

The results of this report confirm that the original BHMmetals, the recently developed BHM mud and TBI are all working to give consistent results. The indices are robust with respect to long-term cycles caused by climate patterns, yet some trends have been detected. Particularly important for the success of the BHM in associating changes in community composition to either heavy metals or mud content is the finding that only rarely do we observe changes in both CAPmetals and CAPmud scores simultaneously. The BHMmetals, BHM mud and TBI all produce complementary information (composition, functionality and resilience), with sensitivity ranges that are also complementary. The TBI also integrates across the interactions between the effect of heavy metals and mud on the macrofaunal communities. For these reasons we recommend using all three indices. We suggest that integrative health scores can be determined for a site.

The BHMmetals is based on a Principal Components Analysis (PCA) axis of total copper, lead and zinc, as this technique deals with the correlation between the metal concentrations. At the time, it was known that the lead concentrations in the seafloor sediments were likely to decrease now that lead was no longer used in petrol. While decreases in lead concentrations are now being observed at some sites, we found no evidence that this change was affecting the ability of the BHMmetals to detect change. An increasing trend in health was observed at a site with decreasing lead concentrations, and multi-year cycles and decreases in health were apparent at other sites that were not demonstrating changes in lead concentrations, or had increases in other metal concentrations. We also found relatively little change as yet in sites where mangrove encroachment is occurring. Thus, there is no need for a special sampling trip to relocate these sites. However, over the longer-term it would be appropriate to relocate these sites to areas lacking mangroves.

Amalgamating health scores across a whole report card area is considered unwise as not all muddy, healthy or contaminated areas have been sampled proportionally. However, the GIS plots we drew of the individual sites and their rankings manage to show how much has been sampled, approximate sample locations and still give information on the overall health. Amalgamating changes over time is less problematic. A report card could simply show a pie chart with segments of red (proportion getting worse), yellow (staying the same), green (getting better). The yellow segment should probably be split into a blue segment (already ranked as good) and an orange segment (staying the same and not good).

Comparison of trends in community composition (BHMmetals) and stormwater contaminants (from Mills et al. 2012) were generally consistent with one another, although slightly more trends were detected for

stormwater contaminants than for community composition. The differences in the number of trends detected may reflect both biological responses (i.e., namely that biological responses may lag behind changes in contaminant chemistry and there may be hysteresis) as well as differing time periods for assessment and trend detection methodologies, which we recommend to be standardised in the future.

Contents

List of figures

List of tables

1.0 Background

In 2001, the ARC held a workshop entitled "Urban Marine Environmental Objectives" to define general quality objectives for urban coastal marine areas in terms of ecology, amenity and health. At the workshop it became apparent that there were no available scientifically defensible assays for, or even definitions of, marine community health. As a result, NIWA was commissioned by the ARC to "define the level of natural variation in marine coastal and estuarine benthic communities in such a way as to enable the development of relevant criteria for health of urban coastal marine areas".

Ultimately, a Benthic Health Model—hereafter referred to as the BHM—was developed with the purpose of providing a tool whereby new observational data of benthic macrofaunal community structure at a given site within the Auckland region could be classified into a category of relative ecosystem health. The first iterations of the BHM were based on qualitative rankings of pollution at a number of sites along a gradient (1 being "healthy" and 5 being "polluted") (Anderson et al. 2002). Eventually, as more data on environmental stressors such as mud (percent of particles <63 μm) and heavy metal contaminants (copper, lead and zinc concentrations) became available, this information was incorporated into the BHM allowing for further refinement and development (Hewitt 2005, Anderson et al. 2006, Hewitt and Ellis 2010).

To date, the BHM has tested as extremely robust and addresses changes in benthic community structure associated with gradients in mud and heavy metal contaminants. Essentially the BHM is a multivariate analysis of macrobenthic community composition (a canonical analysis of principal co-ordinates; CAP). The CAP draws out a gradient in community composition associated with an environmental variable or group of variables of interest. Separate models have been created for two of the key environmental contaminants in Auckland area estuaries: sediment mud content (percent silt+clay, CAPmud) and sediment heavy metal concentrations (copper, lead and zinc; CAPmetals). Thus, although macrofaunal community composition is affected by many different environmental and biotic factors, the models draw out those responses that are only associated with mud or heavy metal concentrations respectively. The models are based on data from 95 intertidal estuarine sites sampled between 2002 and 2005, encompassing mean abundances of more than 100 different soft-sediment macrofaunal taxa. The CAP scores resulting from each BHM model provide, by themselves, a relative rating of ecosystem health, although CAPmetals scores have also been converted into five health groupings (Group 1 being the healthiest, Group 5 the least healthy) and into a green-amber-red reporting system (green being healthy, yellow intermediate, and red unhealthy).

As another step in creating understandable yet scientifically defensible indicators of the ecological integrity of estuarine and coastal areas, a functional traits based index called the TBI was developed (van Houte-Howes and Lohrer 2010) and refined (Lohrer and Rodil 2011). The TBI is now being trialled for use in AC's State of the Environment reporting. This 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, shape, mobility, feeding mode, position in the sediment, sediment reworking behaviour, and type of topographic feature created). 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 TBI 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. Higher numbers of species per functional group probably equates to a greater range of activity types within functions as well. Therefore, the TBI analysis is meaningful with regards to maintaining ecosystem multi-functionality.

Although the TBI is generally less sensitive to mud and heavy metal pollution gradients than the previously developed BHM, the TBI provides more information on whether functional redundancy is changing and whether specific functional traits are being affected. Furthermore, the TBI 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 TBI and BHM may complement each other well by providing a balance of sensitivity, information content and broad general applicability.

1.1 Study Objectives and tasks

The overarching objective of the current study was to test the compatibility, comparability and complementarity of BHM and TBI output and to assess the suitability of these health indicators for State of the Environment reporting. To do this, we organised the work into a number of related tasks. These included:

- Calculation of BHMmetals, BHMmud, and TBI status for all sites where appropriate data was available, including RDP sites (n=95) and various AC ecological monitoring sites (UWH, CWH, Manukau, Mahurangi, Estuarine Monitoring), using data post 2010.
- Investigation of the ability of each index to track temporal changes in ecosystem health at sites that have been sampled more than once.
- Comparisons of index findings to trends in contaminants from the contaminant monitoring programme (Mills and Williamson 2012).
- Investigation of whether recalibration of BHMmetals is required due to the changing relationships between the metal contaminants.
- Investigation of up to 8 sites where mangrove encroachment and channel formation have occurred since the time that the sites were selected to address the following questions: Do differences in benthic community confirm these sites are unsuitable? Is mangrove encroachment and infilling likely to be an increasing issue with upper estuarine sites? How do we sample the health of these areas if the indices are affected by these site characteristics?
- Highlighting the sites where health scores were worse than expected on the basis of CAPmetal and CAPmud scores alone. This exercise may identify sites where further contaminant investigation (e.g., hydrocarbons, sulphide, nutrients) may be warranted.
- Investigation of how regional reporting might complement/overlap with the reporting of individual SOE programmes. Are the results consistent with what we know from our SOE monitoring?
- Investigation of how the three indices interact. In other words, do the three indices provide more information together than on their own? Are they consistent in the health assessment that they provide?
- Reporting the scores for sites sampled recently enough to reflect “current” health status. Presentation of the results in easy to understand tabular and map formats that can be used for general public information including report cards.
- Investigation of the usefulness of the indices within the context of report card areas. For example, is there enough coverage per report card area, can the scores be averaged to provide one score per report card area, etc?

- Consideration of the use of the three indices as regional reporting mechanisms. Is the current coverage of sites enough/representative? Should model development sites be resampled or are there other sites that would be more appropriate to sample to increase coverage?

2.0 Methodology

2.1 Traits Based Index, TBI[1]

Organisms can be categorised taxonomically and also according to characteristics that are likely to reflect ecosystem function (i.e., their feeding mode, degree of motility, position in the sediment column, body size, body shape, capacity to create tubes/pits/mounds, and so on). During 2010 and 2011, an index based on these biological traits was created (van Houte-Howes and Lohrer 2010) and improved (Lohrer and Rodil 2011). The index was based upon seven particular biological traits, representing broad categories relevant to ecosystem function (see Table 2.1).

All seven of the traits used in the index had strong and significant negative responses to both mud and metals (average $r < -0.5$). The seven selected traits 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).

[1] The TBI was originally called the NIWACOOBII, which admittedly was a dippy name. We suggest TBI now.

Table 2-1 A listing of eight broad functional trait categories (left column) and the 32 individual traits 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 TBI 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 TBI calculation (see Lohrer and Rodil 2011 for details). L:W = length to width ratio.

Functional Category	Functional Group	Code
Living position *	Attached	Attached
	Deeper than 2 cm	Deep
	Surface epifauna	Epif
	Top 2 cm	Top
Sediment topography feature created *	Permanent burrow	Burr
	Erect structure / tube	Erect
	Simple hole or pit	Hole
	Mound	Mound
	Trample marks	Trample
	Trough	Trough
Direction of sediment particle movement *	Depth to depth	DD
	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 (L:W 3-10)	Streamlined
	Round/Globulose (L:W 1-3)	Globular
	Worm-shaped (L:W 10-100)	Worm
Body hardness	Soft-bodied	Soft
	Rigid (chitonous skeleton)	Rigid
	Calcified (fully calcified shell)	Calcified

Index values are calculated as follows:

1. The taxonomic richness in each of the 7 trait groups per site are summed (i.e., $Ntaxa_{Top}$ + $Ntaxa_{Erect}$ + $Ntaxa_{SS}$ + $Ntaxa_{Sedentary}$ + $Ntaxa_{Sus}$ + $Ntaxa_{Medium}$+ $Ntaxa_{Worm}$) to produce a quantity called SUM_{actual}.

2. A maximum expected value called SUM_{max} (i.e., a non-polluted reference value) is identified from Table 2 of Lohrer and Rodil (2011). This quantity varies depending on the number of replicate samples used to calculate SUM_{actual} (Table 2.2).

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

4. The TBI formula is **$1 - (SUM_{max} - SUM_{actual}) / SUM_{max}$**, which essentially standardises the index values to fall between 0 and 1. Values near 0 indicate highly degraded sites, and values near 1 indicate the opposite.

Table 2-2 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. This was based on taxa accumulation curves (table is reproduced from Lohrer and Rodil 2011, see for details).

No. of reps	Top	Erect	SS	Sed	Sus	Medium	Worm	SUMmaxn
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

2.2 Benthic Health Model

The original benthic health model (BHMmetals) was developed by Auckland Council (then Auckland Regional Council), Marti Anderson (then Auckland University) and Simon Thrush and Jud Hewitt (NIWA), to determine the health of macrofaunal communities relative to storm-water contaminants. The model was a multivariate analysis of macrobenthic community composition backed by information on total sediment copper, lead and zinc concentrations, extracted from the 500 µm fraction of the sediment (Anderson et al. 2006).

In 2010-2011, a similar model was developed, this time to determine health relative to sediment mud content (BHMmud, Hewitt and Ellis 2011). At the time of the development of this model it was determined that, while there was some crossover between community compositions found in response to high mud and high contaminants, the two effects could still be separated.

Both models are based on the community composition observed at 84 intertidal sites in the Auckland Region between 2002 and 2005. As some sites were sampled more than once, the total number of data points in the model is 95. The sites are within tidal creeks, estuaries or habours, but do not include exposed beaches. They cover a range of contaminant concentrations and mud content. The models use Canonical Analysis of Principal Coordinates (CAP, Anderson and Willis 2003) of square root transformed Bray-Curtis dissimilarities[2]. For the metal model, the concentrations of the three metals have been used in a Principle Component Analysis to create a single axis (PC1) that explains >90% of the variability in contaminant differences between the sites. For the mud model, the % mud content of sediment at the time of sampling is used.

[2] Note that these are based on the average abundance of taxa present in the original 95 sites/times at the taxonomic resolution common to all.

Within the PERMANOVA addon (Anderson et al. 2008) to the Primer E software (Clarke and Gorley 2006), there is the facility to run a CAP and then find out what CAP scores new sites would be allocated. Thus for BHMmetals, the similarity matrix of all sites (the original 95 data points plus however many new sites you wish information on) is opened, and the CAP analysis run using the 95 datapoints of either mud content or PC1, with the “allocate new samples” mode ticked.

3.0 Functional Trait Analysis

3.1 Seasonal patterns

Seasonal patterns in TBI scores were examined using two long term data sets from regional ecological monitoring. TBI scores were calculated at five intertidal sampling sites in Mahurangi Harbour (Hamilton Landing, HL; Jamieson's Bay, JB; Mid Harbour, MH; Te Kapa Inlet, TK; Cowan's Bay, CB) using all available data between July 1994 and October 2006. There was high temporal variability in TBI scores, particularly at JB (Figure 3.1). For example, index scores at JB repeatedly increased and decreased by more than 0.25 in approximately 6 monthly cycles. Scores at JB rose as high as 0.82 in 2006 and dipped as low as 0.36 in 1999. Scores were almost universally higher at JB than they were at the other 4 intertidal sites during this 12 year monitoring period. JB is the most heterogeneous of the sites, having high coarse sand and gravel fractions and generally low but temporally variable mud content values (usually <3%, but rising to 27% in 1996 and >10% in 2001, 2002 and 2006). These factors may explain some of the variability in TBI scores during this time frame.

Scores were the lowest at HL, the muddiest and most homogenous of the sites, although scores at MH and CB appeared to be only slightly higher judging from the full time series. Scores at MH dipped below 0.2 twice during the period (the lowest individual scores recorded, Figure 3.1), with a mean value that hovered around 0.3. Interestingly, the up and down patterns were roughly synchronised at all 5 sites, suggesting that the macrofauna at the five sites may be responding to the same Harbour-wide or seasonal-scale (or greater) drivers.

TBI scores were also calculated at Cape Horn (CH) and Clarkes Beach (CB) in the Manukau Harbour on each sampling occasion. The time series data collected at sites in the Manukau Harbour are among the longest time series data sets that the Auckland Council possesses. Here we consider a 20 year period from December 1987 to December 2007. Site CB has been sampled continuously during this time, while sampling at some of the other sites (e.g., CH) has been phased in and out for various reasons.

As with the Mahurangi data, TBI scores based on all available time series data from the Manukau sites (i.e., 6 sampling occasions per year) were extremely variable (Figure 3.2). Wintertime low scores of <0.3 at CB soared to >0.6 by summertime 6 months later in both 1993 and 1995. The largest change overall was observed 1996-97, when the summertime high score of >0.75 fell to <0.3 by the end of the following winter.

Although it is relatively easy to calculate the TBI as often as six times per year, we would not recommend this, as the marked seasonal variability in TBI scores could be misinterpreted as rapidly vacillating ecosystem health (which we do not believe it is). Instead, we recommend calculating the scores on a once-yearly basis (e.g., every October) and assessing trends over time frames of 5 or more years (see section 4.2). We also recommend viewing TBI scores in relation to stress levels (e.g., Figures 4.3 and 4.4).

Figure 4-1 Time series of TBI scores at five ecological monitoring sites in Mahurangi Harbour. All available data between July 1994 and October 2006 are plotted.

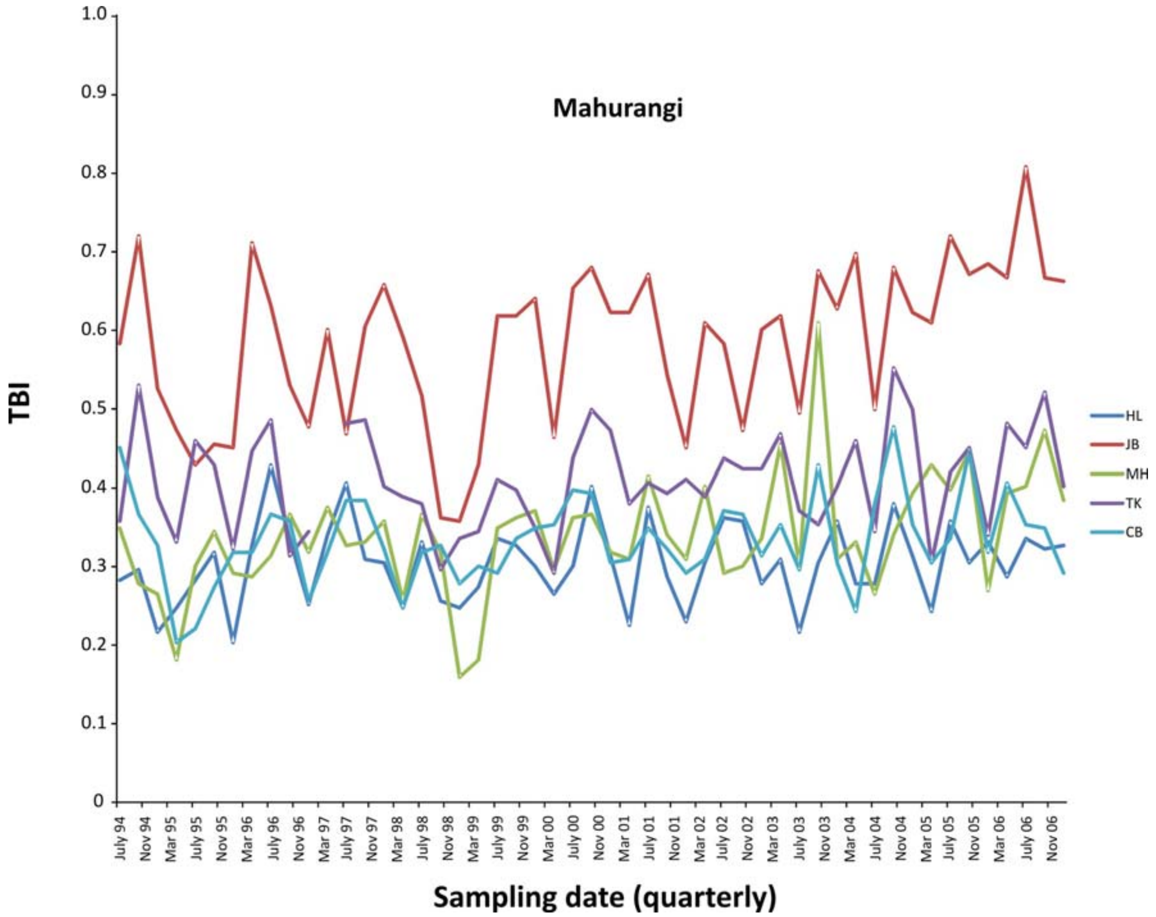
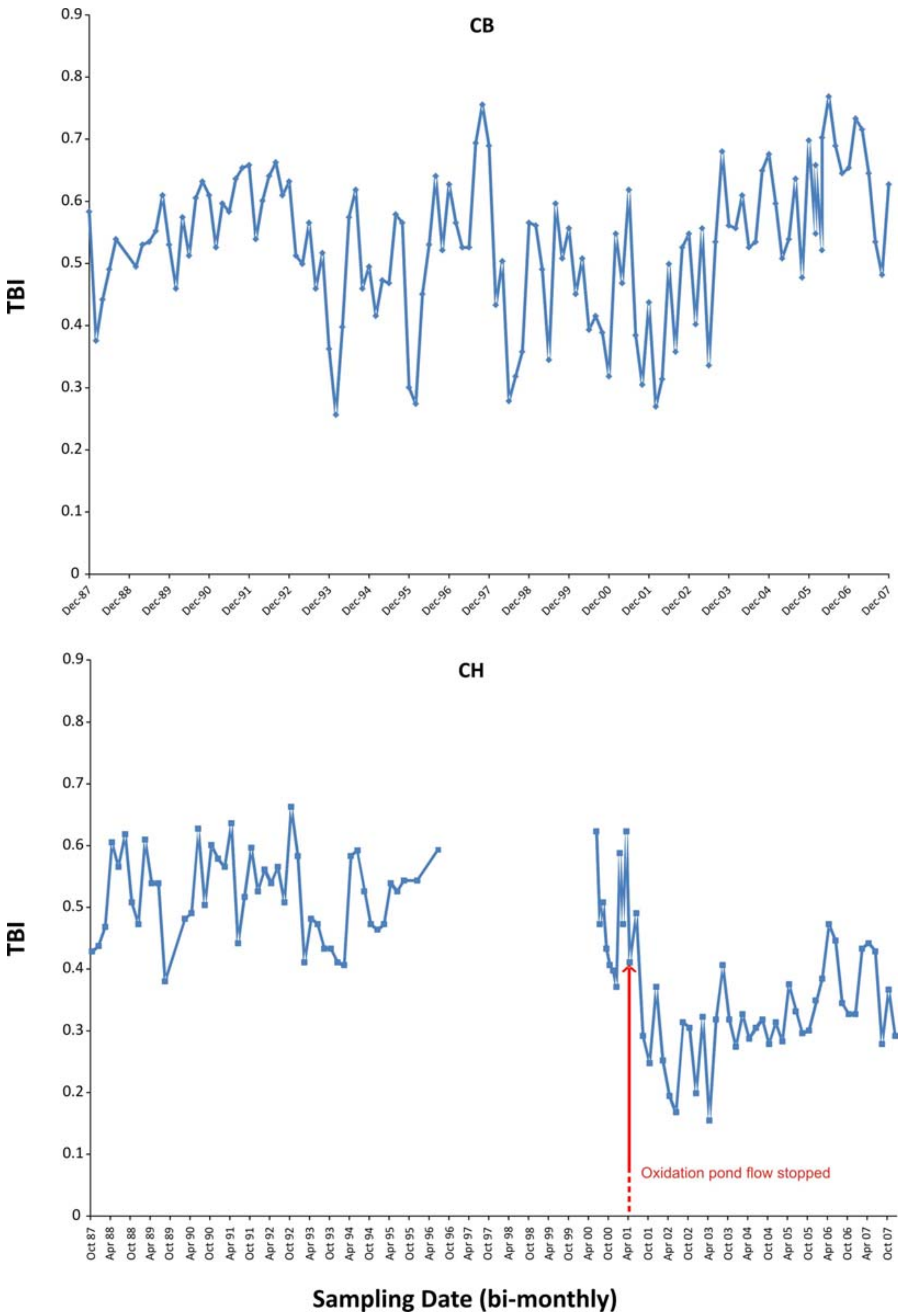


Figure 4-2 Time series of TBI scores at sites CB and CH, from Manukau Harbour. Data between 1987 and 2007 are plotted. Sampling was stopped for a period at CH, and re-started to capture the effects of the cessation of oxidation pond discharge. The discharge was stopped in April of 2001.



3.2 Temporal trends

Statistical trend analysis consisted of Pearson's R or Spearmans ρ correlations, as we wanted to focus on the correlation coefficients rather than the p-values, as p-values for trend analysis are highly sensitive to numbers of samples and our number of samples was low.

3.2.1 Mahurangi Harbour

Temporal trends in TBI scores based on October-only data at all five of the intertidal monitoring sites in Mahurangi Harbour were slightly positive during the 1994 to 2006 period (Figure 3.3), with the upward trend strongest at JB. Importantly, the TBI scores remained, on average, >0.3 at all of the sites for the entire 12 year period. Thus there is no evidence of declining or low levels of functional redundancy at any of the five intertidal monitoring sites in Mahurangi Harbour. Sites JB and TK had average TBI scores >0.4, suggesting that the functional redundancy of the benthic communities at these sites is intermediate or good (i.e., no reason for concern, especially with temporal trends in the positive direction). Sediment mud content appears to have decreased at all five of the Mahurangi Harbour sites during the 1994 to 2006 period (following peaks in mud content during 1996), coincident with the gradual increases in TBI functional redundancy (Figure 3.2). Mud content was also a relatively good predictor of TBI among sites (Figure 3.4).

The most recent TBI scores at JB, MH, TK and HL (2011) were all >0.4 and continue to fit the temporal and mud gradient trends discussed above.

Table 4-1 The most recent available TBI scores for four Mahurangi Harbour sites.

October 2011	TBI
HL	0.42
MH	0.43
TK	0.48
JB	0.83

Figure 4-3 Temporal trends in TBI scores (red lines) and sediment mud content values (blue lines) at five sites in Mahurangi Harbour. For presentation purposes (to fit on the same scale as mud %), all TBI scores were multiplied by 100. Only one TBI score per year (October) is presented, minimising seasonal variability.

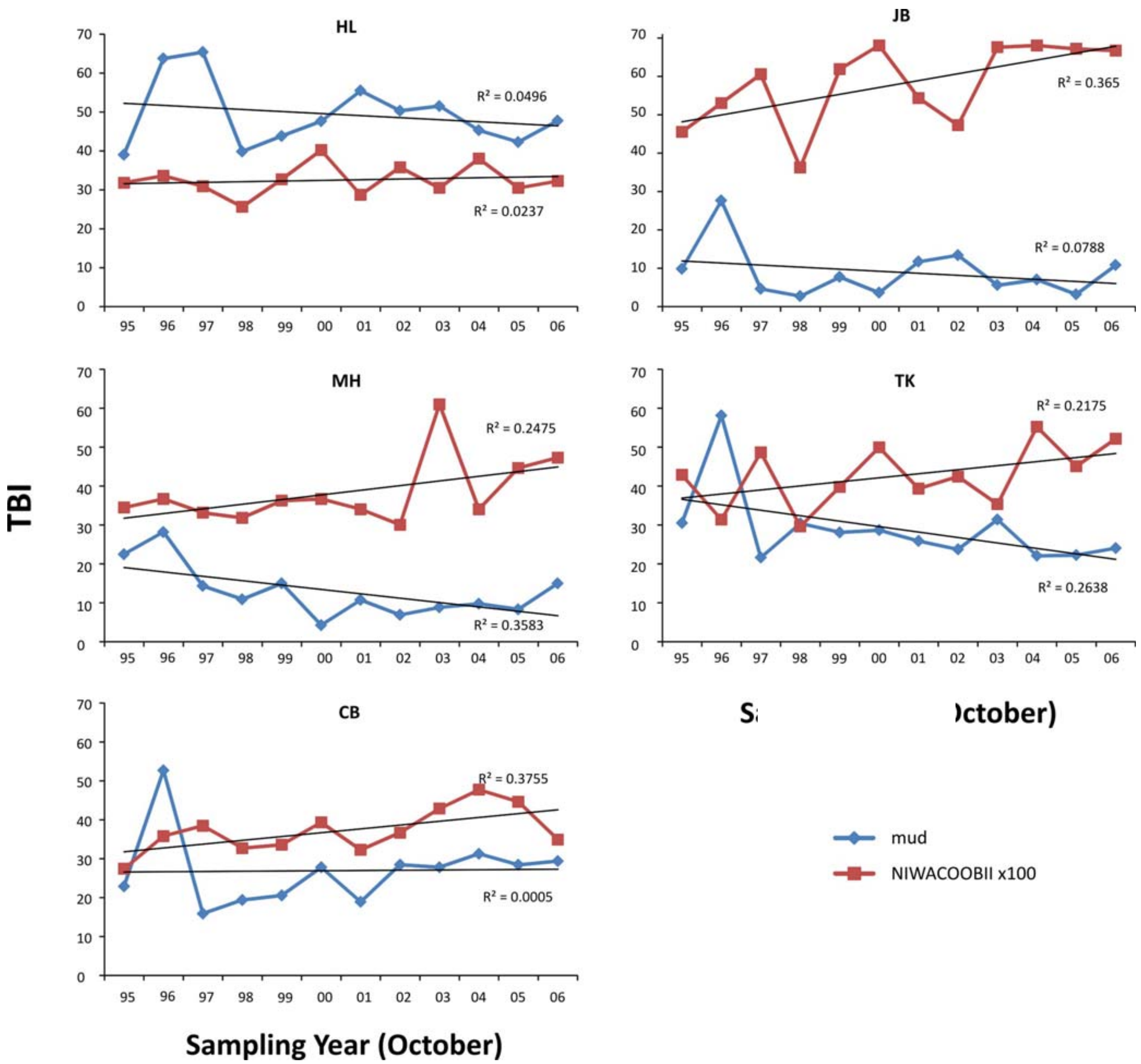
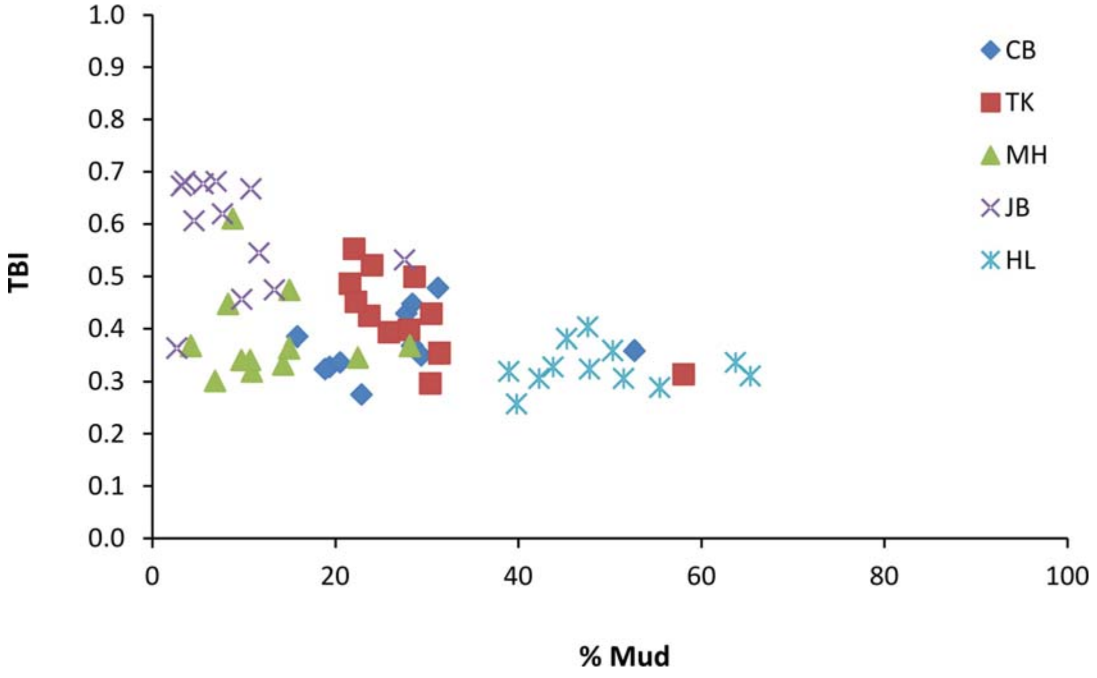
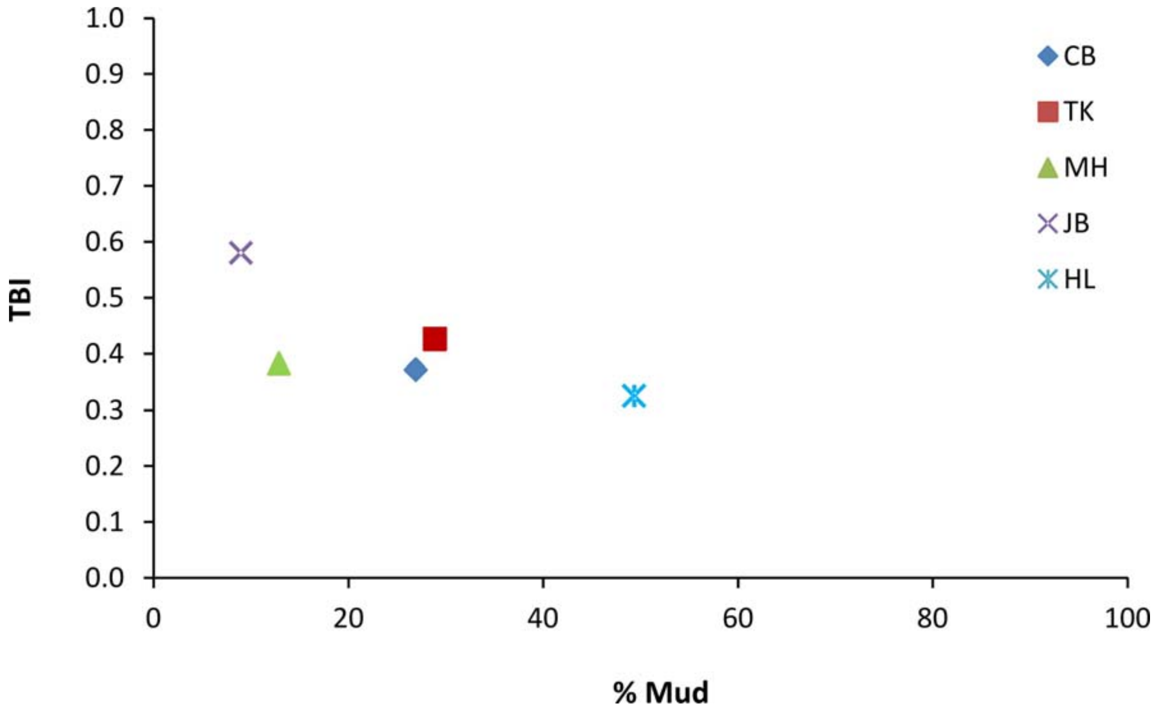


Figure 4-4 TBI scores from each October between 1995 and 2006 plotted versus sediment mud content data from those same sampling periods. The lower panel shows averages for each site over the same period.



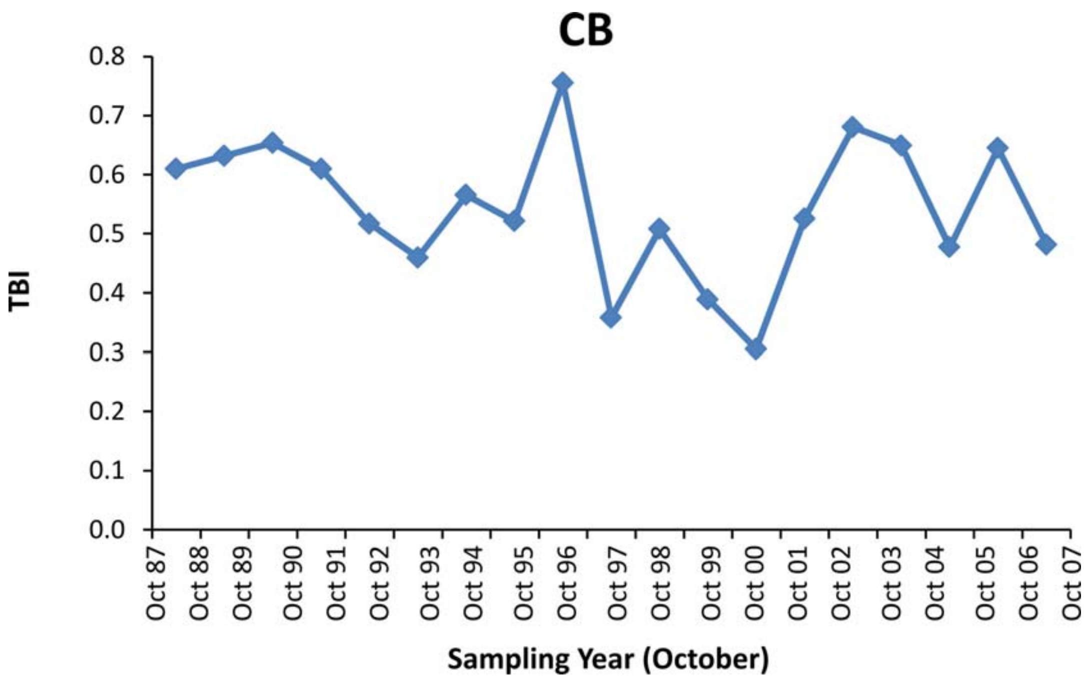


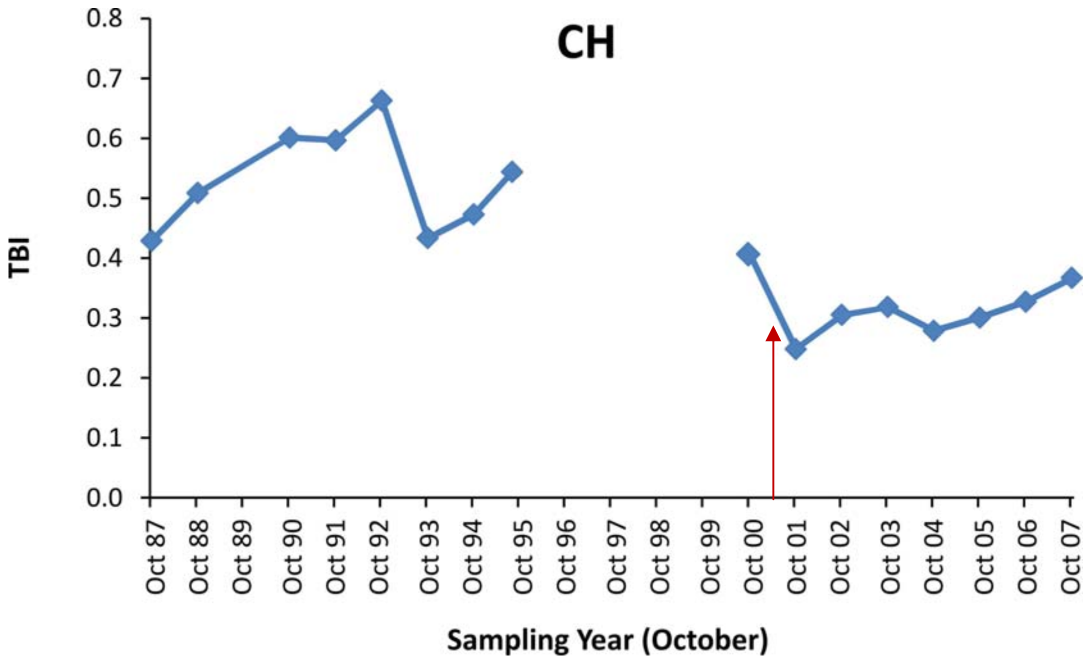
3.2.2 Manukau Harbour

Site CH in the Manukau is an interesting case, as the time series between 1987 and 1996 relates to a period when effluent from a sewage oxidation pond was being discharged 7 km upstream of this site. The time series was reinstated again in 1999 to be able to capture the effects of the cessation of the oxidation pond discharge, which effectively reduced the amount of nutrient and organic matter loading at the site. Discharge from the oxidation ponds was stopped in April 2001. Despite considerable variability in the TBI scores based on the full time series (as at CB), a precipitous drop in TBI scores is evident between April 2001 and May 2003, after which the scores began to slowly rise again. We suggest that the sudden cessation of organic matter loading (although positive for the environment in the long run) represented a disturbance to a community geared towards eutrophic conditions (see also Hewitt and Thrush 2009). The loss of species following the discharge stoppage was tracked by trends in the TBI scores, and there appears to have been some slight recovery since the wintertime lows of 2002 and 2003 (although this may be a general harbour-scale or ENSO-related increase given that the scores at Manukau site CB also rise during this time frame).

Based on October samplings at CH since 2002 Figure 3.5), TBI scores seem to staying above the 0.3 mark for the most part, with no significant downward trend to cause concern. That being said, with an average score of 0.3, the richness of taxa in functional groupings that respond negatively to elevated mud and heavy metals is rather low. Thus, whilst sediment mud content at CH is very low (averaging <2%) and metals are likely to be low (as they were in 2002), increases in either mud or metals would likely impair benthic community health at these sites. Site CB, in contrast, is a heterogeneous site that continues to have very high TBI scores. In combination with the sandy sediments (<10% mud) and low metal concentrations present at this site, it is safe to say that the health status of site CB is good.

Figure 3.5 TBI scores from each October between 1987 and 2007 at sites CB and CH in Manukau Harbour. The break in the CH time series reflects a period when monitoring was stopped. Monitoring at CH restarted prior to April 2001 (red arrow), when the discharge of oxidation pond water 7 km upstream of the site was stopped.





3.2.2 Central Waitemata Harbour

TBI scores were calculated for six Central Waitemata Harbour ecological monitoring sites (Hobsonville, HBV; Henderson Creek, HC; Shoal Bay, ShB; Whau, Reef, and Lower Shoal, LoS) using October data collected between 2000 and 2011. Data from HC and Reef were only available until 2009; data collection at LoS started in 2009. TBI scores have been >0.4 since 2004 at all of the CWH sites and all showed similar significant positive trends in TBI scores dating back to 2001 (Figures 4.6 and 4.7). The sandy sediments, the low metal concentrations, the current scores > 0.4 and the positive trends at HBV, HC, ShB, Whau, and Reef all support the conclusion that the health of these sites is good (no reason for concern). The only possible cause for concern is LoS, which had increasing mud content and decreasing TBI scores between 2010 and 2011; note, however, this is based on just two years of data (Figures 4.6 and 4.8).

Figure 4-6 Time series of TBI scores based on October-only data collected between 2000 and 2011 at 6 sites in the Central Waitemata Harbour.

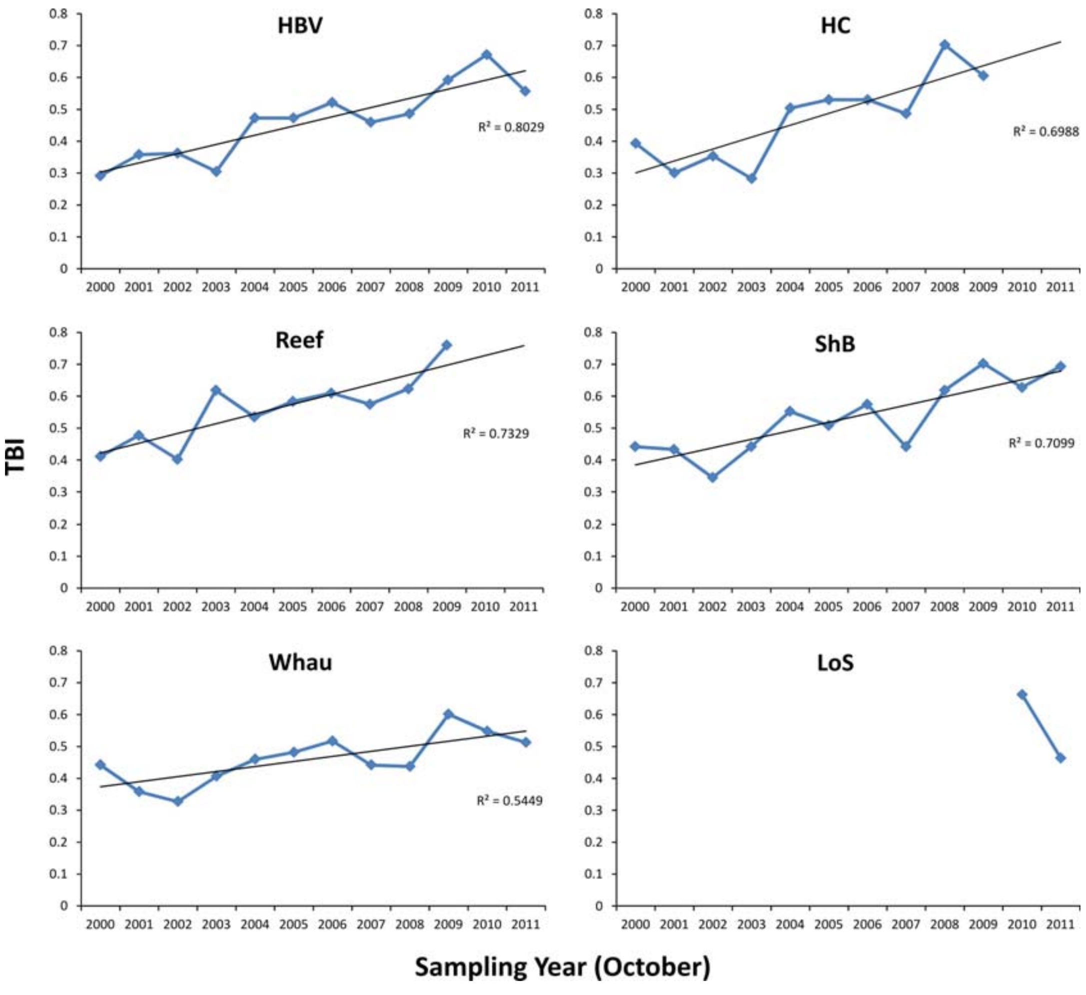


Figure 4-7 TBI scores from each October between 2000 and 2011 at six Central Waitemata Harbour sites plotted versus sediment mud content data. The data are divided by the period in the time series (showing the general rise in TBI scores across the harbour over time).

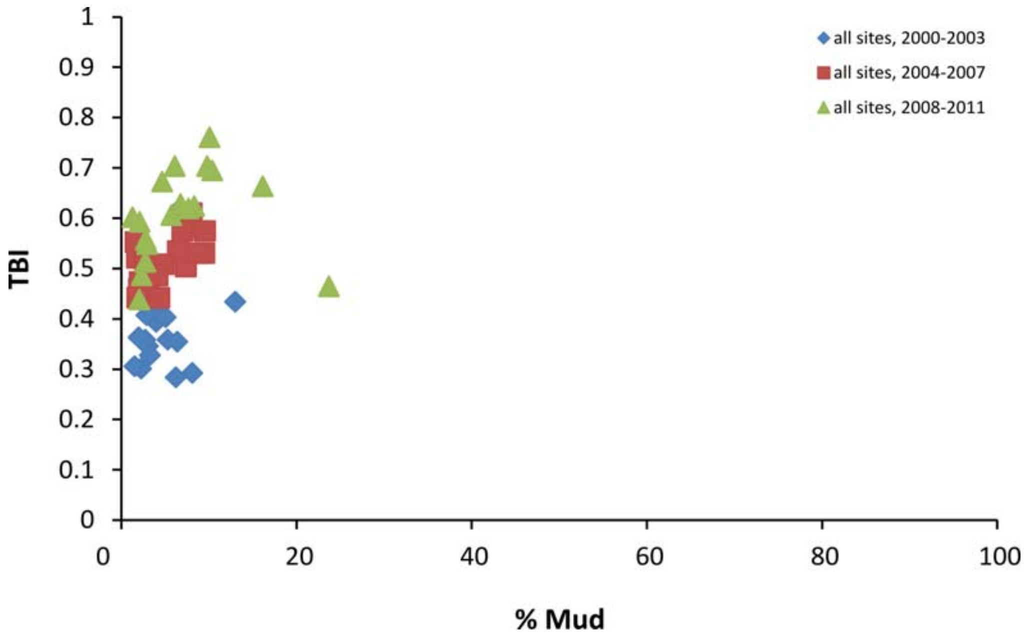
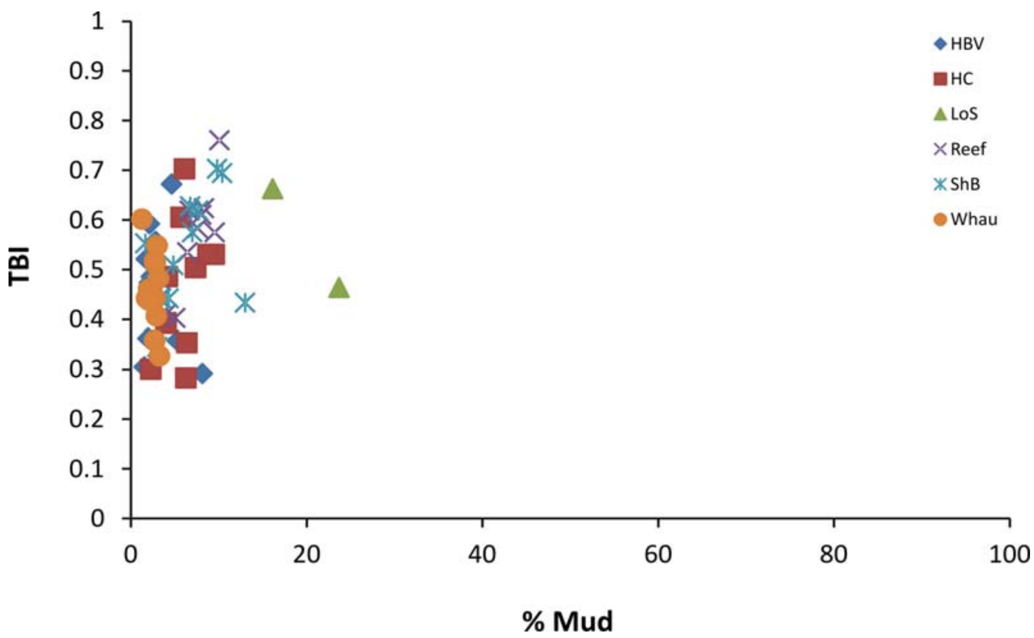


Figure 3.8 TBI scores from each October between 2000 and 2011 at six Central Waitemata Harbour sites plotted versus sediment mud content data. This figure shows the variability in the data at the individual sites.



3.2.3 Upper Waitemata Harbour

The AC monitoring sites in the Upper Waitemata Harbour (UWH) are generally muddier and more contaminated by metals than the ecological monitoring sites in the Central Waitemata Harbour (CWH, discussed above). Therefore, we expected the TBI scores to be generally lower in the UWH, relative to CWH.

Time series from nine UWH sites, with data collected each October between 2005 and 2011, were available for analysis (Figure 3.9). Information on sediment characteristics including mud percentages and heavy metal concentrations were available at all sites (although not sampled every year). For heavy metal concentrations, we used a combined metal index score called PCA1.500, which is calculated from concentrations of Cu, Pb and Zn measured in the <500 µm sediment fraction. PCA1.500 values >0 indicate sites with relatively high metal contaminant concentrations; negative values indicate cleaner sites (with sites down near -2.5 amongst the very cleanest).

The temporal trends in TBI scores were generally weak, with low slope or low r2 values or both (Table 3.1). Exceptions to this were Brighams Creek (BRIG) and Main Upper (MainU), which had negative slopes with r2 > 0.4.

Table 4-1 Trends in time series data for Upper Waitemata Harbour sites, based on November data 2005-2011.

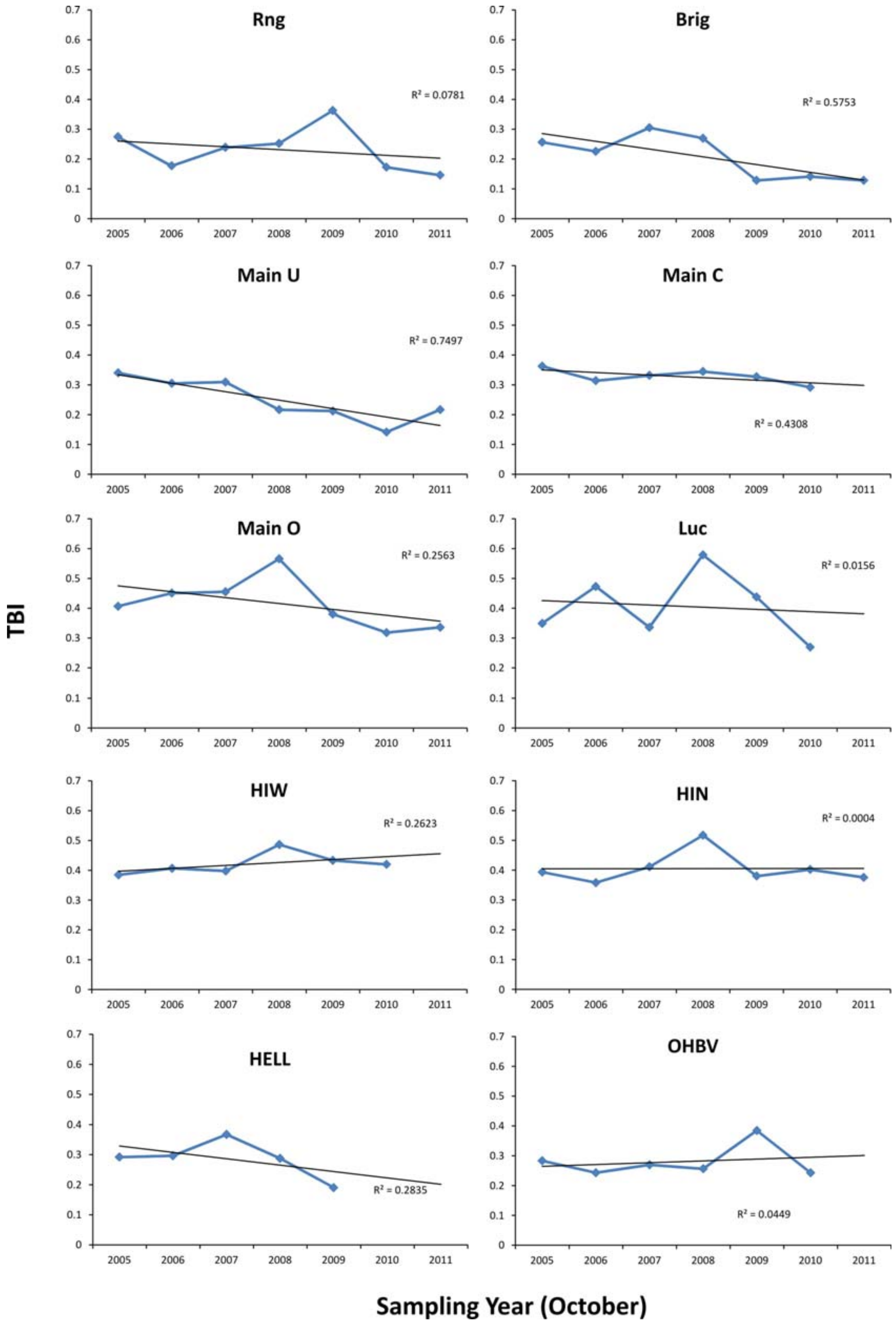
Site	P-value	r^2	slope	Pearson's R
Brighams Creek (BRIG)	0.0481	0.5753	-0.03	-0.76
Herald Island North (HIN)	0.9889	0	<.001	0.00
Herald Island Waiarohia (HIW)	0.2989	0.2623	0.01	0.51
Lucas Creek (LUC)	0.8134	0.0156	-0.01	-0.12
Main Central (MainC)	0.1568	0.4308	-0.01	-0.66
Main Outer (MainO)	0.2463	0.2563	-0.02	-.51
Main Upper (MainU)	0.0118	0.7497	-0.03	-0.87
Opposite Hobsonville (OHBV)	0.6870	0.0449	0.01	0.21
Rangitopuni Creek (RNG)	0.5439	0.0781	-0.01	-0.28

The TBI scores at BRIG and MainU indicate areas of poor ecological health (scores trending below 0.2, average mud content of 85-90%, average PCA1.500 scores >0.6). MainC is a concern, as (1) its average TBI score over the last seven years was just above the 0.3 threshold value, (2) it showed a significant downward trend during this period, (3) it had a moderately high mud content value, averaging 24.4%, and (4) it had a moderately high sediment heavy metals score, average PCA1.500 = 0.33.

LUC was another site with borderline-high mud and metals results (mud 32%, PCA1.500 = 0.31), although this site was less concerning due to a higher average TBI score (0.41) and no significant temporal change during the period. Thus, although the site is moderately muddy and contaminated, it appears to have higher levels of functional redundancy than MainC (i.e., higher richness of macrofaunal in groupings susceptible to mud and metals).

The site with the poorest overall condition was RNG (average TBI score of 0.23 during the 2005-2011 period). Site OHBV was also poor. MainO was intermediate, in that it was a relatively muddy site (16%) but not overly contaminated (PCA1.500=-0.10), with a good TBI score (0.41). Sites HIN and HIW appeared to be in the best condition (TBI scores > 0.4, mud content <10%, PCA1.500<-0.5).

Figure 4-9 Time series of TBI scores based on October-only data collected between 2005 and 2011 at 10 sites in the Upper Waitemata Harbour.

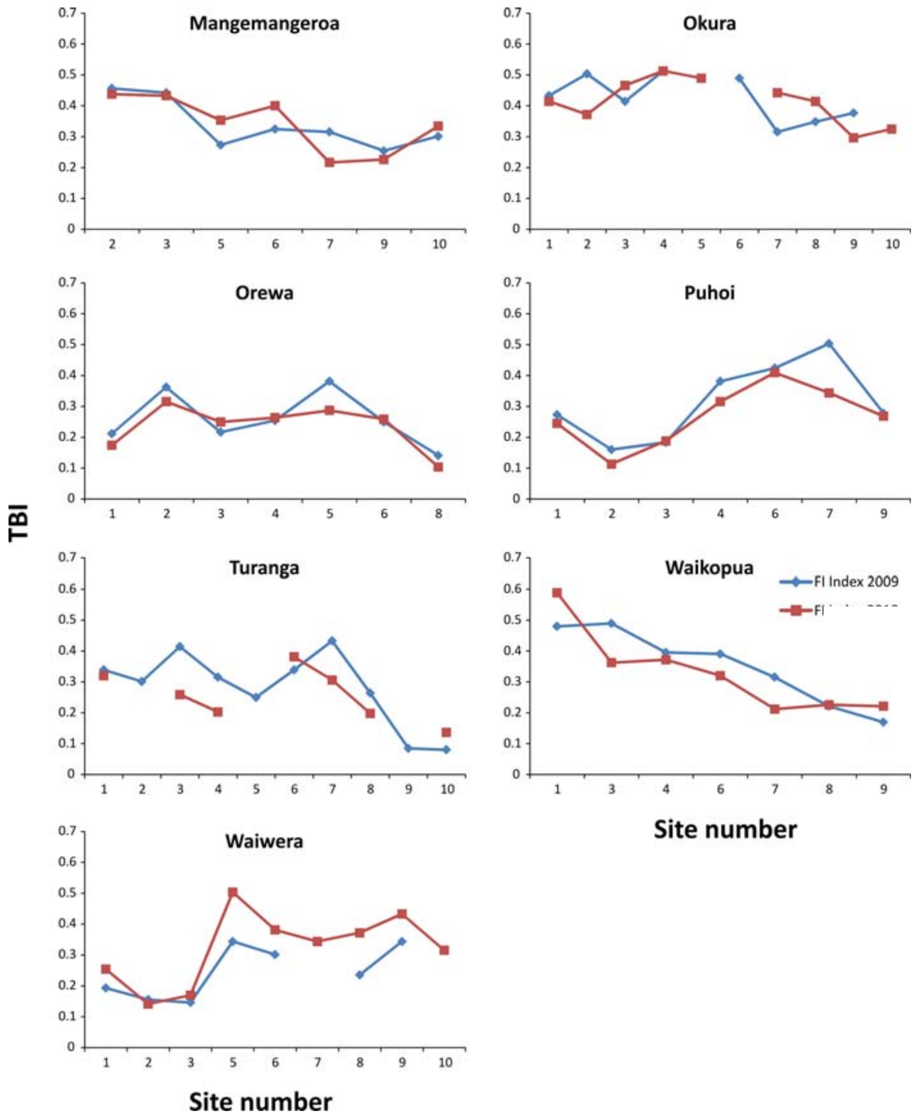


3.3 Consistency over time in spatial patterns

TBI scores were calculated using October data from 2009 and 2010, collected from 7 to 10 sites per estuary, in 7 different estuaries (Mangemangeroa, Turanga, Orewa, Puhoi, Waikopua, Waiwera and Okura). These data were suitable for examining spatial changes running along head-to-mouth estuarine gradients, as well as the consistency of spatial patterns in two consecutive years.

Taking the second issue first, there was remarkable consistency in the spatial patterns between years in all 7 of the estuaries sampled (Figure 3.10).

Figure 4-10 TBI scores at multiple sites within the seven Estuarine Monitoring Programme estuaries in two consecutive years (blue: 2009, red: 2010). Sites are listed along the x-axis of each individual panel; low site numbers are near estuary mouths.



The spatial patterns in each of the seven estuaries were mainly related to changes in grain size. The lowest TBI scores were generally associated with fine sediments (>25% mud). The highest TBI scores were associated with sandy sediments with 2.5 to 10% mud. Interestingly, sites with almost no mud at all tended to have slightly lower TBI again, which is likely related to physical disturbance by waves and currents.

A pattern of lower than expected TBI scores was observed at sites with <2.5% mud and a higher percentage of coarse sand than mud. This pattern was reported previously in a wave-swept part of the Wairoa Embayment in eastern Tamaki Straight (Site A-int; Lohrer et al. Wairoa report). Sediments in hydrodynamically active areas tend to have limited numbers of individuals and species due to the physical disturbance and transport of the sediments by waves and currents. The low TBI scores at hydrodynamically active sites likely reflect this, rather than negative response to mud or metals. In the seven estuaries considered here, the hydrodynamically active sites with low TBI scores appeared to be at the edges of swiftly moving tidal channels rather than in locations exposed to waves. Fitting this description are Sites 1 and 3 at Orewa; Sites 1 and 4 at Puhoi; Sites 1 and 3 at Turanga; Sites 3, 6 and 8 at Waiwera).

All seven estuaries had TBI scores ≤0.3 at two or more sites. Okura had the highest average scores across sites during both years, with only one value <0.3 in each year (at different sites each year), indicating that this is the healthiest of the Estuarine Monitoring programme sites with respect to the TBI score. Sediment mud content data supports this conclusion also: Okura had the lowest average mud content across sites and between years (9.8%).

Mangemangeroa and Turanga were the muddiest two estuaries on average. The muddiest upstream sites in each estuary (Sites 8-10, averaging almost 40% mud) had TBI scores averaging 0.2. Sites 9 and 10 at Turanga had scores <0.1 in 2009, indicating very low functional richness and redundancy.

A decline in TBI scores with increasing mud content was observable each year in all seven of the individual estuaries, particularly when the hydrodynamically active sites (<2.5% mud and coarse fraction > mud fraction) were removed from the analysis. (Figure 3.11).

The humped shaped relationship between sediment mud content and TBI scores is clearly evident in Figure 3.12, which shows data from all sites and times where we possess both mud content and TBI information. The relationship rises between approximately 0 and 2.5% mud, peaks between 2.5 and 10%, drops between 10 and 25% and descends more slowly between 25 and 100%.

Figure 3.11 TBI scores at multiple sites within the seven Estuarine Monitoring Programme estuaries in two consecutive years (blue: 2009, red: 2010). Scores are plotted versus sediment mud content for each estuary, with significant gradients in sediment mud content present in each one. Note the drop in TBI scores with increasing sediment mud content, as well as the anomalously low TBI scores when there is almost no mud present at all (<2.5%).

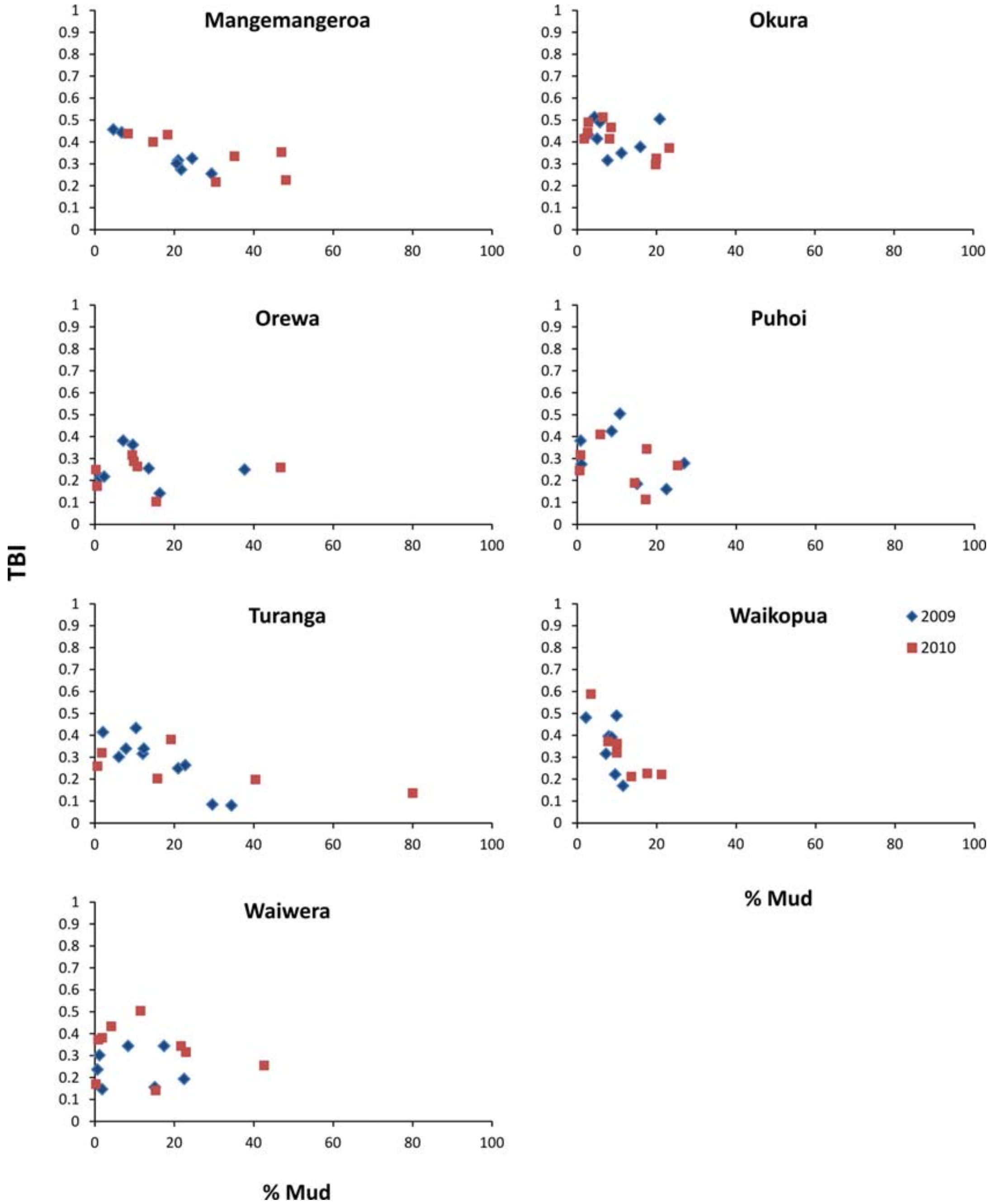
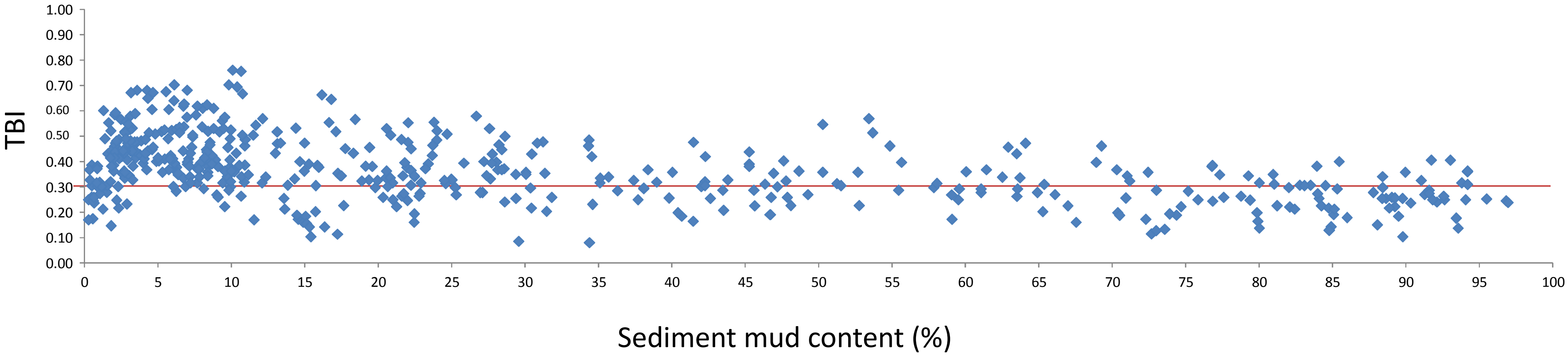


Figure 3.12 TBI scores versus sediment mud content using all available data (multiple estuaries, multiple years, etc.).



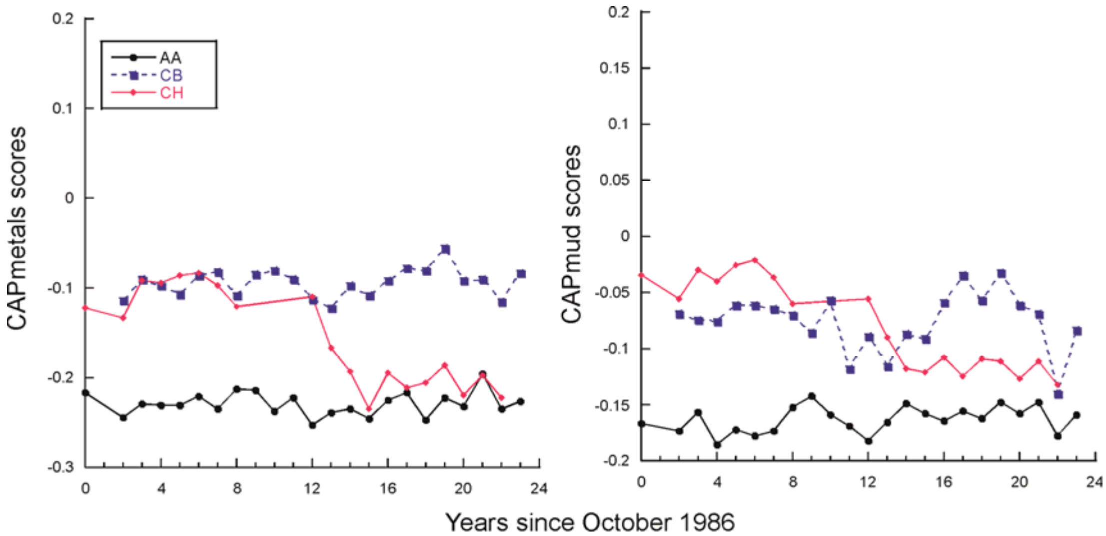
4.0 Benthic Health Model Analysis

4.1 Temporal patterns

The potential for multi-year cycles to adversely affect the ability of the BHM to accurately associate change in macrofaunal community composition to contaminants was initially tested by comparing samples collected at the same sites at two different times (2 – 3 years apart). This analysis suggested that temporal variability on this time scale was not of concern (Anderson et al. 2002). Here we test sensitivity over a longer time period (23 years) using the Manukau Ecological Monitoring Programme at the three long-term monitored sites.

Two of the three sites (Auckland Airport and Clarkes Beach) both show very little variation over the 23 years in CAPmetals scores, with slightly more for CAPmud (Figure 4.1). However, the third site, (Cape Horn), located 7 km away from Mangere on a sand flat in the middle of the harbour) exhibits a decline in CAPmetals and CAPmud scores associated with changes in community composition occurring post the removal of the oxidation ponds. There is some indication from the Manukau Ecological Monitoring Programme that mud content has declined slightly at this site.

Figure 4-1 Changes over time in CAPmetals and CAPmud scores observed at sites Auckland Airport, Clarkes Beach and Cape Horn.



4.2 Observed trends over time

Trend analysis was conducted for all sites for which routine analysis of stormwater contamination is conducted (Mills and Williamson 2012; sites from the SOE, RDP and UWH monitoring programmes). Similar to that report, analyses were only conducted on sites with more than 3 sampling points. This results in some sites being analysed by Mills and Williamson (2012) for contaminant trends but not for ecological changes as they are not sampled with the same frequency. Conversely, we were able to include 2011 samples in our analysis. The statistical trend analysis consisted of Pearson's R or Spearmans □ correlations, as we wanted to focus on the correlation coefficients rather than the p-values, as p-values for trend analysis are highly sensitive to numbers of samples and our number of samples was low.

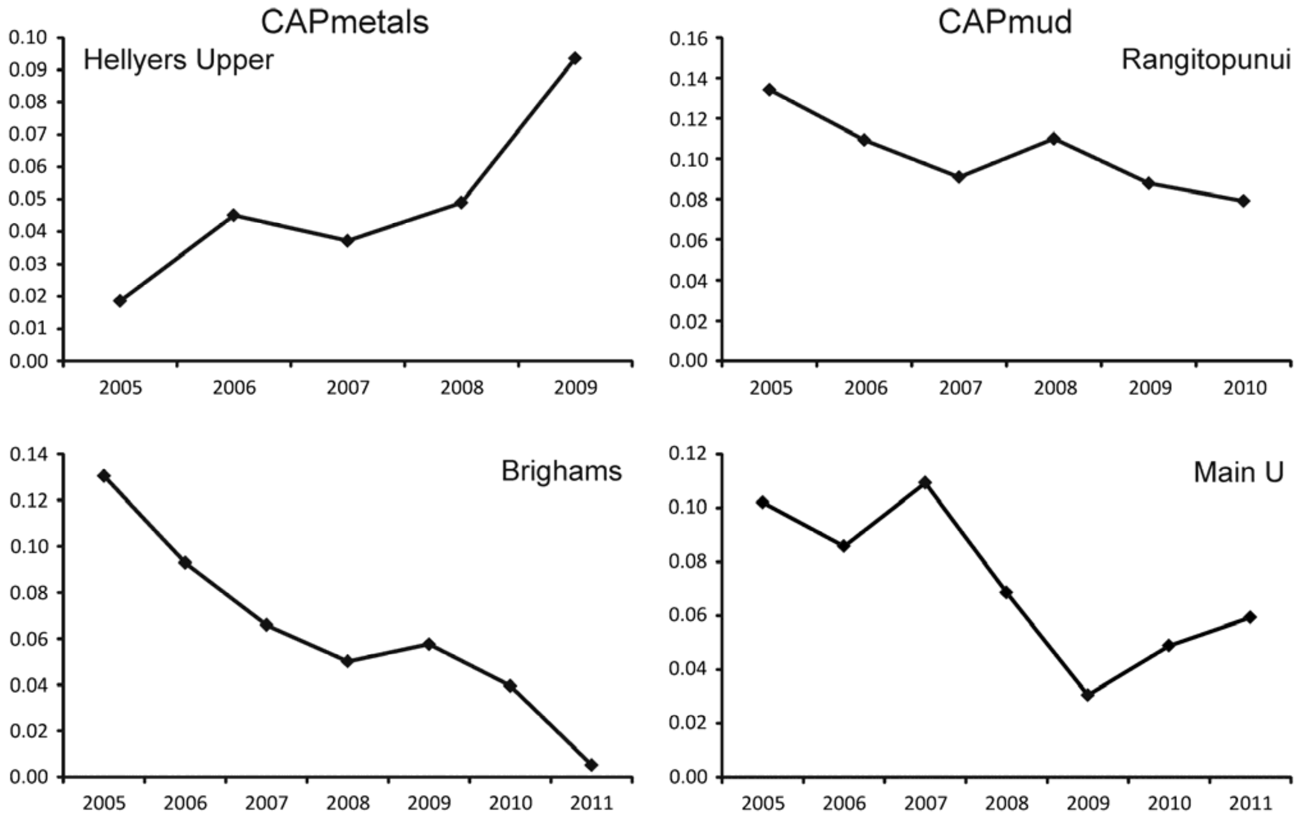
Table 4-1 Sites with correlation coefficients of CAPmetals and CAPmud greater than 0.7 where trends do not seem likely to be part of multi-year cycles or to be driven by a single point at either end of the time series. Increasing health has negative correlations, decreasing health has positive correlations.

Site	CAPmetals	CAPmud
a) RDP sites		
Panmure	-1.00	
Newmarket	-0.83	
Whau Lower	-0.92	
Awatea		-0.84
Whau Entrance		-0.91
b) UWH sites		
Rangitopunui (UWH)		-0.87
Main U (UWH)		-0.75
Lucas Upper (UWH)	-0.86	
Hellyers Upper (UWH)	0.90	
Herald Island North (UWH)		0.86
Brighams (UWH)	-0.96	-0.89

Of the RDP sites, only three significant trends for CAPmetals were observed; all were increases in health (Table 4.1). The trends in community composition observed at Whau Lower and Newmarket are consistent with the decreasing trends in total copper and lead concentrations found at these sites. However, for Panmure only small positive increases in copper and zinc have been detected. For CAPmud scores there were a number of correlations > 0.7 driven by either a single high or low point at the beginning of the time series (4 and 7 correlations respectively). Only two changes over the time series were observed, increases in health at Awatea and Whau Entrance respectively (Table 4.1).

More high correlations that were related to consistent changes in community composition were observed at the Upper Waitemata Harbour sites, probably due to the longer time series (Table 4.1). Trends of increasing health were observed at Rangitopunui, Main U and Brighams related to mud and at Lucas Upper and Brighams related to contaminants. Trends of decreasing health were observed at Herald Island North related to mud and at Hellyers Upper related to contaminants (Figure 4.2).

Figure 4-2 Changes over time in CAPmetals scores observed at Hellyers Upper and Brigham sites, and CAPmud scores observed at Rangitopunui and Main U sites, all within the Upper Waitemata Harbour. Increasing health has negative correlations, decreasing health has positive correlations.



All of the changes observed were of <0.02 per year, that is, the site score changed by < 5% of the total range in CAP scores used by the model. Thus, only sites on the edges of groups changed their rankings. However, over the monitored period this resulted in changes in CAPmetals scores varying from 7% (Panmure Table 4.2) to 18% (Brighams), and for CAPmud, from 2.5% (HIN) to 16% (Brigham). Changes of 25% represent movement from the mid-point of one group to the mid-point of the next group.

Table 4-2 Magnitude of changes over time observed at sites in Table 4.1.

Site	Model	change/yr	%change/yr	%change over monitored period
Brighams (UWH)	metals	-0.0175	0.035	0.17
Hellyers Upper (UWH)	metals	0.0150	0.030	0.15
Lucas Upper (UWH)	metals	-0.0150	0.030	0.15
Panmure	metals	0.0061	0.012	0.07
Newmarket	metals	-0.0250	0.050	0.03
Whau Lower	metals	-0.0098	0.020	0.12
Awatea	mud	-0.0014	0.004	0.02
Brighams (UWH)	mud	-0.0125	0.031	0.15
Herald Island North (UWH)	mud	0.0023	0.006	0.03
Main U (UWH)	mud	-0.0100	0.025	0.12
Rangitoponui (UWH)	mud	-0.0092	0.023	0.11
Whau Entrance	mud	-0.0052	0.013	0.10

4.3 Mangrove encroachment

Three sites were highlighted as a major concern due to encroachment by mangroves into the site (Henderson Upper, Newmarket and Oakley), and another 3 highlighted as a potential for concern (Meola Inner, Motions and Whau Wairau). These sites were all sampled in 2005, 2007, 2009 and 2011 (although some were also sampled in 2002). Four other sites were also sampled at the same time scale (Whau Lower, Middlemore, Meola Reef, Mangere and Ann's Creek). CAPmetals and CAPmud scores were calculated for these sites, and comparisons made over time. Only 2 sites exhibited a consistent trend over time (Newmarket and Whau Lower); both of these were decreases in CAPmetals scores (increases in health) and consistent with the decreasing trends in total copper and lead concentrations observed at these sites. Three other sites exhibited a higher value at the start of the time sequence compared to those observed since (Motions, Middlemore and Mangare for both CAPmetals and CAPmud). All other sites exhibited non-directional variation over their time sequences and the difference between the start and 2011 could be either positive or negative. Differences in the degree of change from start to 2011 were similar for the CAPmetals regardless of degree of mangrove encroachment and < 10% (Table 5.3). Changes in CAPmud scores were slightly higher for the sites with major encroachment (11%, cf 6.0 and 7.7). However, such a small change will not have yet compromised the monitoring programme.

Table 4-3 Summary of changes in CAPmetals and CAPmud scores over time as a percentage of the range of scores derived from the full model (minimum health to maximum health, 0.5 for CAPmetals and 0.4 for CAPmud). Summary is across sites of major concern for mangrove encroachment, minor concern and no concern monitored on odd years from 2005). Trends and times when the first sampling point was markedly different from other times are not included in the summary.

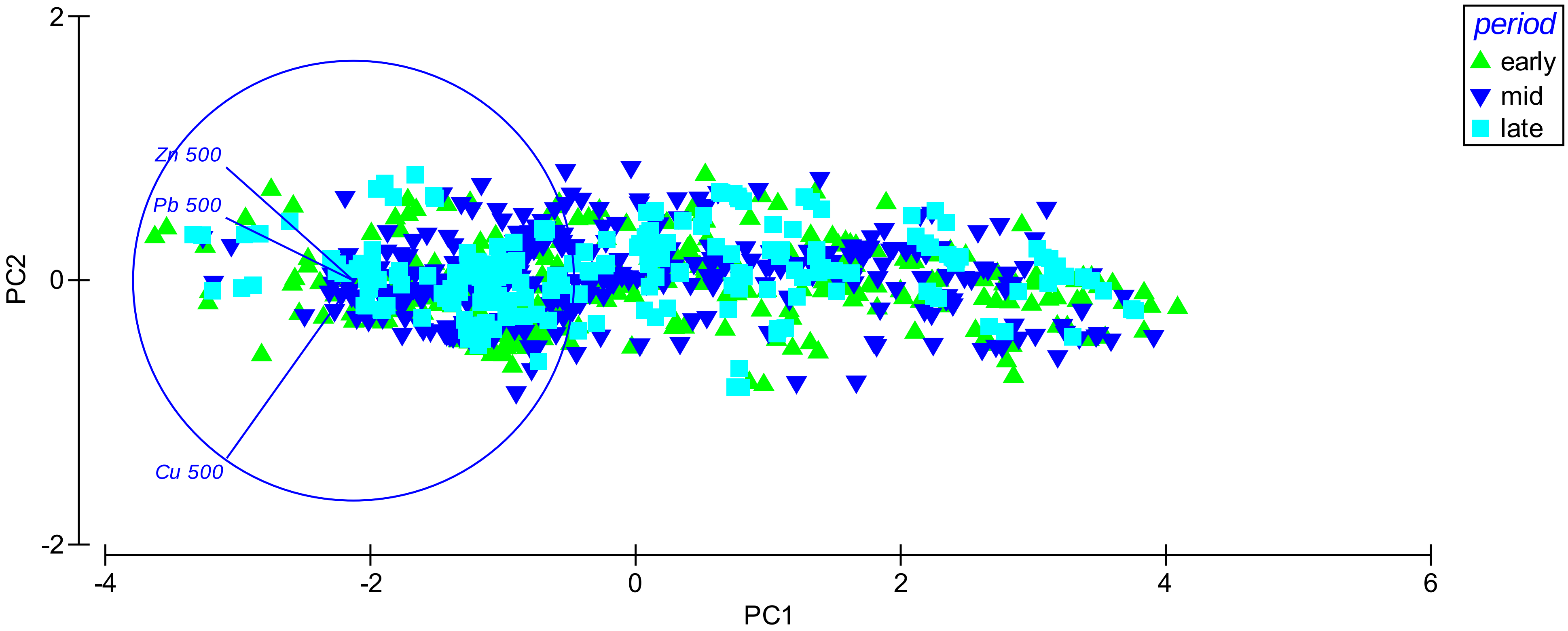
	CAPmetals	CAPmud
Major	6.0	11.0
Minor	8.1	6.0
None	6.6	7.7

4.4 Changes in relative contributions of heavy metals

The BHMmetals is based on a PCA axis of total copper, lead and zinc, as this technique deals with the correlation between the metal concentrations. However, at the time, it was known that the lead concentrations in the seafloor sediments were likely to decrease now that lead was no longer used in petrol. Mills and Williamson (2012) document decreases in lead concentrations at a number of sites. The likelihood of this affecting the BHM was assessed by conducting a PCA on loge(x+1) transformed total copper, lead and zinc concentrations observed at all sites since 2002. In the ordination plot of this analysis, data belonging to the years 2002 – 2005 (containing the samples from which the BHM was developed) was colour coded differentially to those data from 2009 (Figure 4.3). The first axis of the PCA explained over 90% of the variability (as did the one used in the BHM model) and there was no differentiation between sites from before 2006 or after 2008. Thus while changes in the relative contributions of the different heavy metals may be occurring, this is not yet affecting the ability of the BHM model to assess changes relative to overall contaminant concentrations.

We recommend that in future reports this analysis continues to be performed.

Figure 4-3 Principal component analysis of $\log_e(x+1)$ transformed total copper, lead and zinc concentrations observed at all sites since 2002. The first axis of the PCA explains over 90% of the variability. Early = data collected between 2002-2005, mid = data collected between 2006 – 2008, and late = data collected between 2009 – 2010.



5.0 Reporting Considerations

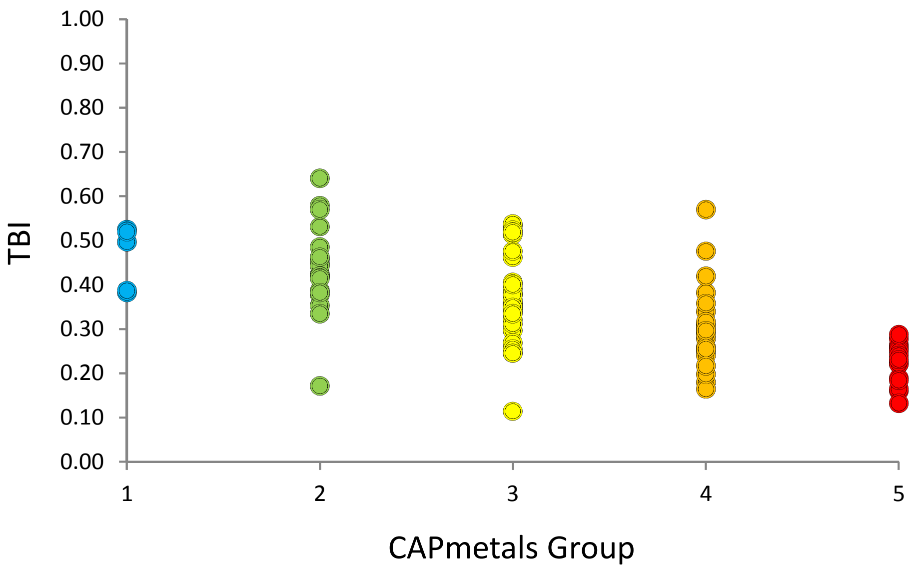
5.1 Comparisons between TBI and BHM results

The data that forms the basis of the BHM models were collected between 2002 and 2005. This is the best data set for making inter-index comparisons as it contains reliable mud, heavy metal and macrofaunal community information from 84-95 sites across a large cross-section of intertidal estuarine habitat types in the Auckland region.

CAP scores from the mud and metals models vary continuously between roughly -0.2 and 0.2, with the lowest scores indicating the healthiest sites. The continuous CAP scores can be converted into categories of health using cut-off values, as was done for the original BHMmetals (Anderson et al. 2006). CAP metal scores < -0.164 indicate "Group 1" sites (i.e., healthiest), CAP metal scores > 0.1 indicate "Group 5" sites (i.e., least healthy); the CAP scores dividing Groups 2-3 and 3-4 are -0.0667 and 0.0234, respectively. In the figures presented in this section, we have colour coded the sites based on the CAP metal groupings, with blue indicating Group 1 (good health), red indicating Group 5 (unhealthy), and green-yellow-orange representing Groups 2-4. In this way, we were able to simultaneously visualise how the BHMmetals rankings compared to the rankings of other health indices (BHM mud and TBI).

Figure 5.1 shows TBI scores relative to the five BHM groupings based on the CAP metal scores. All of the Group 1 sites, and all but one of the Group 2 sites (collectively shown in green), had TBI scores >0.30. In contrast, none of the Group 5 sites (red) had TBI scores >0.30. Thus, 0.3 has potential as a TBI cut-off value that can be used to separate healthy and unhealthy sites. However, sites in BHM Groups 3 and 4 had highly variable TBI scores, some <0.2 and some >0.5, creating some ambiguity that needed to be addressed.

Figure 5-1 Plot of TBI scores versus the CAP metals groups. Data come from the original 95 RDP sites. The colours used to indicate CAP metals groups are used again on subsequent figures.



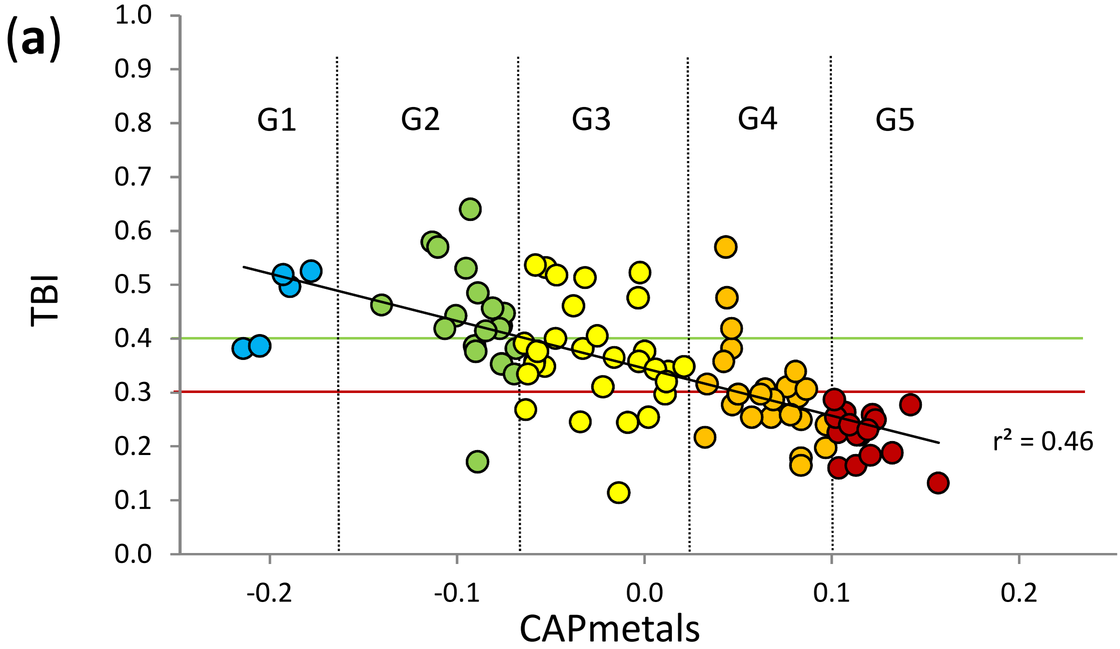
TBI scores were correlated with both types of CAP scores (Figure 5.2, r2 = 0.46 for CAP metals, r2 = 0.48 for CAP mud). The relationship between TBI scores and CAP metals scores was tighter at the upper end of the range (i.e., at relatively contaminated sites, CAP metals > 0.05, Figure 5.2a), whereas the relationship between TBI and CAP mud scores was tighter at the lower end of the range (i.e., at relatively sandy sites, CAP mud < -0.03, Figure 5.2b). TBI scores were noticeably higher when the mud content was less than 10%, with scores ranging between 0.3 and 0.64 and showing up as mostly "green" (no "reds") (Figure 5.3). Beyond this mud threshold, changes in the TBI scores were more subtle, although beyond 25% mud, there were only two "green" sites and TBI scores were commonly in the 0.1 to 0.3 range. Beyond 60% mud, TBI scores never exceeded 0.4 and were predominantly less than 0.3.

Interestingly, the "red" sites, which were defined by a relatively narrow band of CAP metals scores (>0.1, Figure 5.3), had a broader range of CAP mud scores (0.0 to 0.2) and occurred across a large mud gradient (20% to 100%). So it appears that changes in macrofaunal functional redundancy represent an interaction between sediment mud and metal content. The first major decline is driven by mud as it increases between 10 and 25%. The second decline appears to be driven by heavy metal contamination, occurring once sites with > 25% mud become sufficiently contaminated by heavy metals, PCA1.500 > 0.3.

The TBI is not as sensitive an indicator of the response of macrofaunal communities to small changes in mud and heavy metals, with the BHMmetals and BHM mud demonstrating more consistent separation between sites at low levels of stress. However, the ability of the TBI to represent the mud-metal interaction is very useful. We know that the response of macrofaunal species to sediment mud and heavy metals content is not a simple additive response (Thrush et al. 2008), although Hewitt and Ellis (2010) found the amount of variation in community composition explained by the interactive effect of mud and metals to be relatively low. Regardless, having another index that incorporates this interactive effect allows for a more robust measure of overall health.

With respect to this, trends in the TBI which occur above 0.4 and below 0.3 may be considered uninformative, and only trends approaching or crossing these boundaries are important.

Figure 5-2 Plot of TBI scores versus (a) CAP metals scores, and (b) CAP mud scores. The cut-off values used to differentiate the BHM groups are shown with dashed vertical lines. The colours of the dots refer to CAP metals groups. Data come from the original 95 RDP sites.



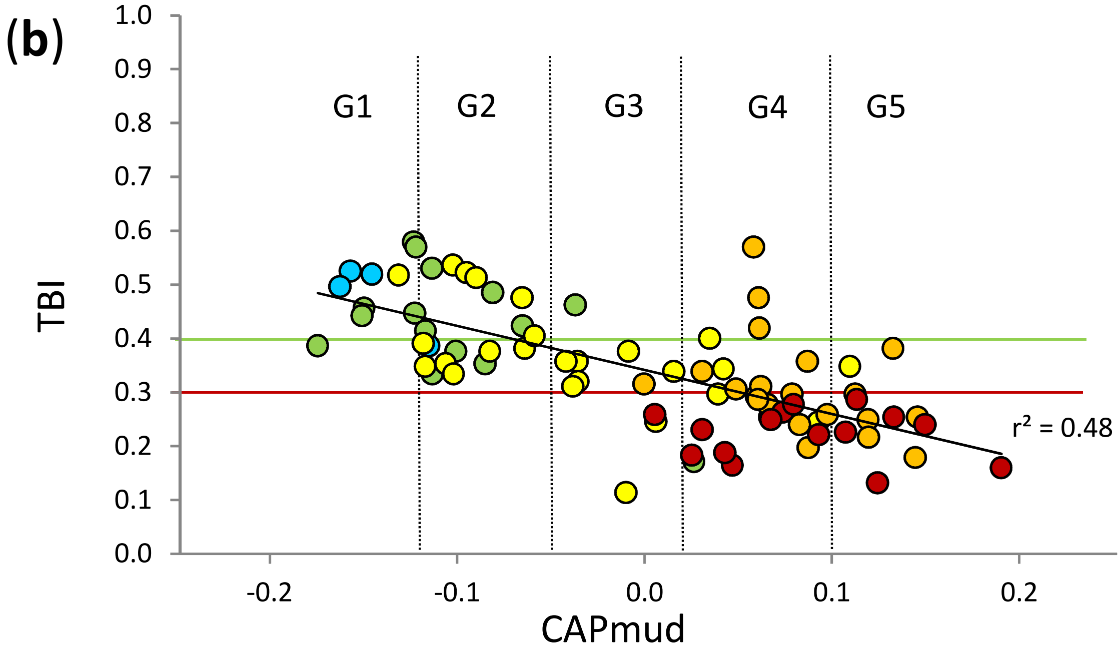
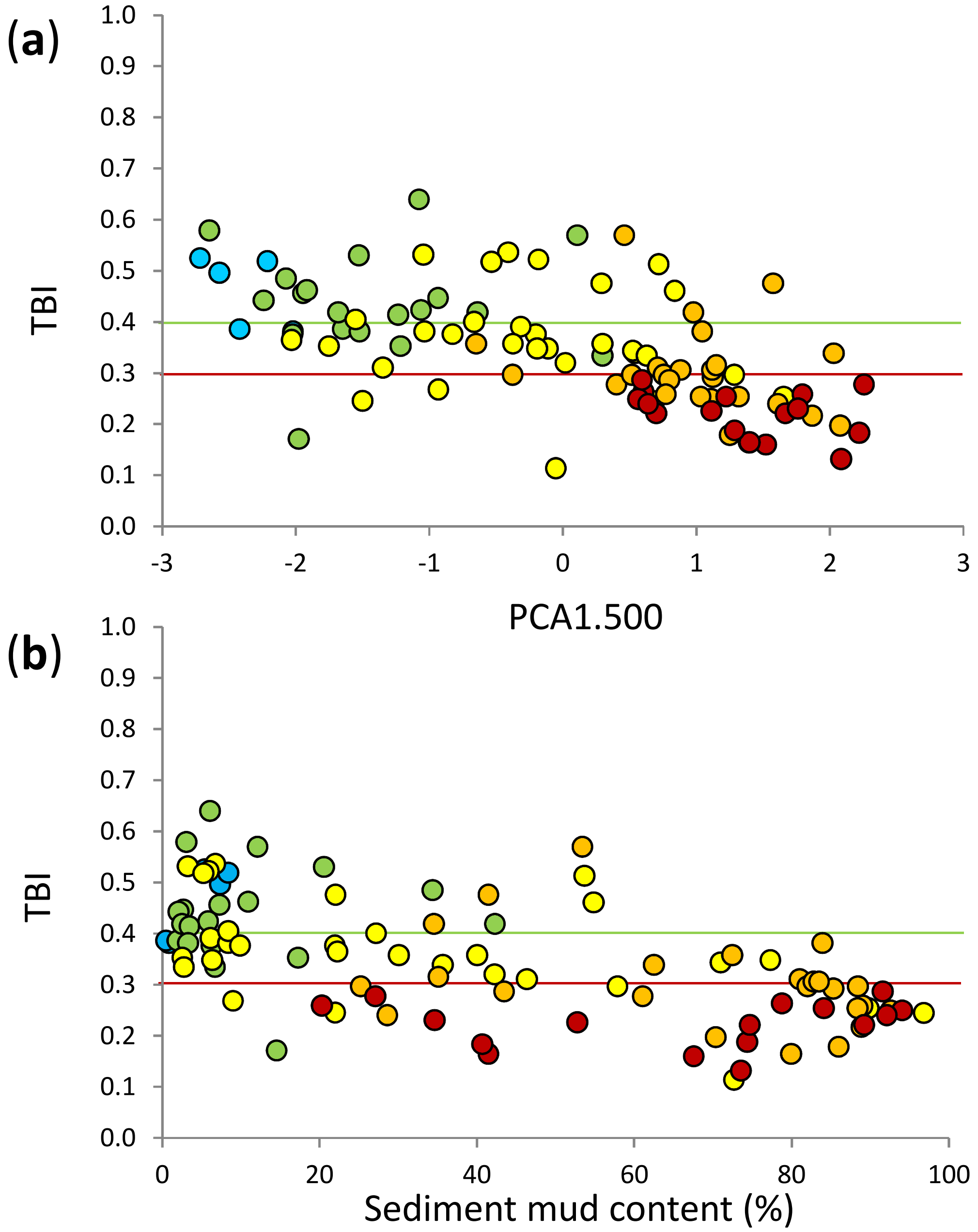


Figure 5-3 Plot of TBI scores versus (a) sediment heavy metals PCA1.500 and (b) sediment mud content. The colours of the dots refer to CAP metals groups. Data come from the original 95 RDP sites.



5.2 Combining results for reporting

Based on the analyses above, we have created a methodology for using the scores from all three methods to judge benthic health.

CAPmud scores were separated into five equal interval groups along the CAPmud axis, and the TBI scores were separated into three groups as per the section above (>0.4, 0.3-0.4 and <0.3). The group values for the CAPmetals, CAPmud and TBI were then converted to range from 0 to 1 with 1 being "poor" (Table 5.1).

Table 5-1 Conversion of CAPmetals, CAPmud and TBI scores into values for producing report cards.

Group	CAPmetals		CAPmud		TBI	
	Cutoff	value	Cutoff	value	Cutoff	value
1	-0.164	0.2	-0.12	0.2	0.4	0.33
2	-0.0667	0.4	-0.05	0.4	0.3	0.67
3	0.0234	0.6	0.02	0.6		1.0
4	0.10	0.8	0.10	0.8		
5		1.0		1.0		

Average health values were then constructed in the following way:

1. If the CAPmud score allocated the site to Mud group 1 then Health was calculated as the average CAPmetals and CAPmud group scores (as we have noted the TBI does not work well when mud content is extremely low).
2. If the CAPmetals score allocated the site to Group 4 or 5, then Health was equal to the TBI group score (reflecting the remaining level of functional redundancy present in these strongly metal-affected communities).
3. Otherwise, health was the average of the CAPmetals, CAPmud and TBI group scores.
4. Recoding these scores as:

    a. $\leq$ 0.2 " extremely good"

    b. 0.2 – 0.4 inclusive "good"

    c. 0.4 – 0.6 exclusive "moderate"

    d. 0.6 – 0.8 exclusive "poor"

    e. $\geq$ 0.8 "unhealthy with low resilience."

Table 5.2 shows the resulting health scores for the original BHM data.

Table 5-2 Proposed health scores for the original BHM data. CAPmetals, CAPmud and TBI groups are given for comparison, where red = "unhealthy", orange = "poor", yellow = "moderate", green = "good", and blue = "extremely good".

Site	PCA1.500	MUD%	CAPmetal	CAPmud	TBI	Metal Grp	Mud Grp	TBI Grp	Health
Anns Creek	1.52	67.56	0.10	0.19	0.16	5	5	3	1.00
Anns Creek	1.25	86.01	0.08	0.14	0.18	4	5	3	1.00
Auckland Airport	-2.02	0.90	-0.21		0.38	1	1	2	0.20
Awatea Rd	0.54	35.70	0.01	0.02	0.34	3	3	2	0.62
Bengazi	-0.20	21.95	0.00	-0.01	0.38	3	3	2	0.62
Bowden Rd	1.29	57.86	0.01	0.04	0.30	3	4	2	0.69
Brigham	0.70	89.25	0.12	0.09	0.22	5	4	3	1.00
Cape Horn	-2.42	0.50	-0.21	-0.12	0.39	1	2	2	0.42
Chelsea	-1.06	5.88	-0.08	-0.07	0.42	2	2	1	0.38
Clarkes Beach	-2.65	3.11	-0.11	-0.12	0.58	2	1	1	0.30
Coxes, Waitemata	-0.93	2.72	-0.07	-0.12	0.45	2	1	1	0.30
Glendowie	-1.94	7.31	-0.08	-0.15	0.46	2	1	1	0.30
Hellyers	0.46	53.43	0.04	0.06	0.57	4	4	1	0.33
Hellyers outer	0.40	61.08	0.05	0.07	0.28	4	4	3	1.00
Henderson Entrance	-0.11	6.39	-0.05	-0.12	0.35	3	2	2	0.56
Henderson Entrance	-0.41	6.78	-0.06	-0.10	0.54	3	2	1	0.44
Henderson Lower	1.11	92.64	0.08	0.12	0.25	4	5	3	1.00
Henderson Upper	1.40	41.47	0.11	0.05	0.16	5	4	3	1.00
Herald Island North	-0.31	6.15	-0.06	-0.12	0.39	3	2	2	0.56
Hillsborough	-0.37	40.07	0.00	-0.04	0.36	3	3	2	0.62
Hobson - Tohunga	-1.04	8.43	-0.03	-0.06	0.38	3	2	2	0.56
Hobsonville	-1.65	1.99	-0.09	-0.17	0.39	2	1	2	0.30
Hobsonville	-2.24	2.12	-0.10	-0.15	0.44	2	1	1	0.30
Kaipatiki	1.13	85.33	0.08	0.06	0.29	4	4	3	1.00
Lower Shoal Bay	-1.55	8.43	-0.03	-0.06	0.40	3	2	1	0.44

Site	PCA1.500	MUD%	CAPmetal	CAPmud	TBI	Metal Grp	Mud Grp	TBI Grp	Health
Lucus outer	0.30	30.08	-0.06	-0.04	0.36	3	3	2	0.62
Lucus Te Wharau	0.71	81.04	0.08	0.06	0.31	4	4	2	0.67
Lucus Upper	0.60	78.78	0.11	0.07	0.26	5	4	3	1.00
Mangemangeroa B	-2.02	6.29	-0.09	-0.10	0.38	2	2	2	0.49
Mangere Cemetery	1.03	88.66	0.07	0.15	0.25	4	5	3	1.00
Harania Creek	0.76	82.03	0.05	0.08	0.30	4	4	2	0.67
Kiwi Esplanade	0.53	71.02	0.01	0.04	0.34	3	4	2	0.69
Meola Inner	2.03	62.53	0.08	0.03	0.34	4	4	2	0.67
Meola Inner	1.61	28.64	0.10	0.08	0.24	4	4	3	1.00
Meola Outer	-1.75	2.58	-0.06	-0.11	0.35	3	2	2	0.56
Meola Reef	-0.18	5.95	0.00	-0.09	0.52	3	2	1	0.44
Meola Reef	0.29	22.05	0.00	-0.07	0.48	3	2	1	0.44
Meola West	0.72	53.70	-0.03	-0.09	0.51	3	2	1	0.44
Middlemore	1.11	52.77	0.10	0.11	0.23	5	5	3	1.00
Middlemore	1.29	74.38	0.13	0.04	0.19	5	4	3	1.00
Motions	1.80	20.34	0.12	0.01	0.26	5	3	3	1.00
Motions	2.26	27.11	0.14	0.08	0.28	5	4	3	1.00
Motions East	-0.53	5.27	-0.05	-0.13	0.52	3	1	1	0.40
Ngataringa Bay	1.05	83.95	0.05	0.13	0.38	4	5	2	0.67
Oakley	1.65	89.77	0.00	0.07	0.25	3	4	3	0.80
Okura D	-2.72	5.44	-0.18	-0.16	0.43	1	1	1	0.20
Orewa F	-2.07	34.37	-0.09	-0.08	0.40	2	2	1	0.38
Orewa G	-2.57	7.38	-0.19	-0.16	0.41	1	1	1	0.20
Otahuhu Creek	1.22	84.11	0.10	0.13	0.25	5	5	3	1.00
Out Main UWH	-0.83	9.91	-0.06	-0.08	0.38	3	2	2	0.56
Pakuranga mid	0.98	34.56	0.05	0.06	0.42	4	4	1	0.33
Panmure	1.12	83.52	0.06	0.05	0.31	4	4	2	0.67
Paremoremo	0.57	94.07	0.12	0.07	0.25	5	4	3	1.00
Paremoremo upper	0.69	96.84	-0.01	0.09	0.24	3	4	3	0.80
Pollen Island	0.02	42.26	0.01	-0.04	0.32	3	3	2	0.62
Princess St	0.80	43.49	0.07	0.06	0.29	4	4	3	1.00
Puhinui	-0.65	72.45	0.04	0.09	0.36	4	4	2	0.67

Site	PCA1.500	MUD%	CAPmetal	CAPmud	TBI	Metal Grp	Mud Grp	TBI Grp	Health
Puhinui, Entrance	-1.35	46.36	-0.02	-0.04	0.31	3	3	2	0.62
Puhoi F	-2.21	8.43	-0.19	-0.15	0.43	1	1	1	0.20
Puhoi H	-1.50	22.01	-0.03	0.01	0.20	3	3	3	0.73
Pukaki	-0.19	77.33	0.02	0.11	0.35	3	5	2	0.76
Purewa	1.15	35.13	0.03	0.00	0.32	4	3	2	0.67
Rangitopuni	0.64	92.13	0.11	0.15	0.24	5	5	3	1.00
Rangitopuni UWH	0.59	91.59	0.10	0.11	0.29	5	5	3	1.00
Shoal Bay, Hillcrest	0.77	89.05	0.08	0.10	0.26	4	4	3	1.00
Shoal Bay, Upper	-1.23	3.49	-0.08	-0.12	0.41	2	2	1	0.38
Turanga G	-1.53	20.59	-0.10	-0.11	0.44	2	2	1	0.38
Turanga J	-0.05	72.69	-0.01	-0.01	0.09	3	3	3	0.73
Upper main UWH	0.52	88.43	0.06	0.11	0.30	4	5	2	0.67
Victoria Ave	-1.21	17.28	-0.08	-0.09	0.35	2	2	2	0.49
Waiwera E	-1.92	10.93	-0.14	-0.04	0.38	2	3	2	0.56
Waiwera J	-1.98	14.57	-0.09	0.03	0.14	2	4	3	0.73
Weiti	-0.66	27.21	-0.05	0.03	0.40	3	4	1	0.58
Whakataka	0.30	6.75	-0.07	-0.11	0.33	2	2	2	0.49
Whakataka	0.11	12.15	-0.11	-0.12	0.57	2	1	1	0.30
Whau East	1.58	41.48	0.04	0.06	0.48	4	4	1	0.33
Whau Entrance	0.63	2.75	-0.06	-0.10	0.33	3	2	2	0.56
Whau Upper	1.76	34.63	0.12	0.03	0.23	5	4	3	1.00
Whau Upper	2.08	70.39	0.10	0.09	0.20	4	4	3	1.00
Whau Wairau	2.22	40.68	0.12	0.03	0.18	5	4	3	1.00
Whau Wairau	2.09	73.58	0.16	0.12	0.13	5	5	3	1.00
Whau West	1.87	88.87	0.03	0.12	0.22	4	5	3	1.00

The sensitivity of the results in Table 5.2 to the TBI group scores was tested by:

- Replacing the group 1 score of 0.33 with 0.16 (i.e., halving it).
- Not performing step 2 above.

Making changes to the TBI group score or the decision tree only made small changes to the resultant health scores, i.e., moving a site from scoring in the top of group b to being just within group c.

It is also possible to derive “health” scores for heavy metals and mud, separately. For mud, steps 1 and 3 are used, with the Group metals scores being left out. For heavy metal contamination, steps 2 and 3 are used, with the Group mud scores being left out.

We also investigated the possibility of using the actual CAPmetals, CAPmud and TBI scores, rather than group scores. This also worked well, but again few differences were observed between the final rankings gained by the two methods.

5.3 Creating scores for report card areas

Trying to amalgamate health scores within a report card area, we immediately ran into 2 major problems.

- Tidal creek areas in the Manukau and Central Waitemata always had lower scores than the sand flat areas. Initially we thought to separate these out into tidal creek vs harbour sites, but this results in the same problems noted in Anderson et al. (2006) and Van Houte-Hawes and Lohrer (2010): where exactly does the division occur?
- Not everywhere has been sampled, and even where the number of sites within a report card area is high, the overall structure of sampling does not necessarily reflect the areal proportions. That is, there is not necessarily the same proportion of sites in contaminated areas vs non-contaminated as the proportion of area that is contaminated vs non-contaminated. Similarly, within an estuary, the muddiest place may have been sampled- but in another one the sampling location may be in a less muddy area, even though the range of sediment mud content may be similar in both. While this problem may be able to be countered for mud by conducting highly structured sampling, it would be impossible to do this for metal contaminants, without a prohibitively costly sampling programme.

If amalgamations were to occur, it would still be necessary to quantify uncertainty by, for example, ranking the number of sites per unit area, thus putting yet another number on the report card.

Finally, amalgamation raises the question of what to express. Should it be the average value, or the best or the worst? For example, both the Central Waitemata and the Manukau areas hold both the healthiest and the least healthy sites (Appendix 1). All of these have drawbacks and while using an average seems inherently satisfying, smearing the results in this way is bound to draw attention from people living beside problem areas who see the average health as not reflecting what they observe.

The GIS maps we drew of the individual sites and their rankings manage to show how much has been sampled, approximate sample locations and still give information on the overall health (Figures 6.4 – 6.7).

Amalgamating changes over time is less problematic. A report card could simply show a pie chart with segments of red (proportion getting worse), yellow (staying the same), and green (getting better). The yellow segment should probably be split into a blue segment (already ranked as good) and an orange segment (staying the same and not good).

The GIS plots of the report card areas clearly demonstrate the uneven coverage of sampling sites within the different areas (see also Table 5.3). Admittedly, the focus is on sampling in areas where problems from storm-water contaminants are expects. Other sites are located in areas where State of the Environment monitoring has been established for other reasons, e.g., Manukau, Mahurangi and Kaipara harbours and the major estuaries along the East Coast. However, there are some obvious gaps, e.g., Bethels Beach estuary, Muriwai, Piha, Wairoa, Kawakawa Bay, Huruhi Bay and Okahuiti Creek on Waiheke and Matakana estuary.

Table 5-3 Report areas with the number of intertidal sites sampled.

Area	Number of sites	Number of estuaries or tidal creeks
Kaipara	6	0
Wellsford/Warkworth	14	2
Mahurangi	6	0
Hibiscus Coast	21	3
East Coast Bays	0	0
Waiheke North	0	0
Tamaki Strait	21	3
Tamaki Estuary	6	2
Manukau	5 - 8	1
West Coast	0	0
Central Waitemata	23	5 (all)
Upper Waitemata	13	6 (all)

Figure 5-4 Health scores post 2010 at all sites for which data are available. Health scores reflect the combination of CAP mud, CAP metals and TBI indices (see section 6.2 on page 42). Health scores are indicated by colour (blue: extremely good, green: good, yellow: moderate, orange: poor, red: unhealthy with low resilience).

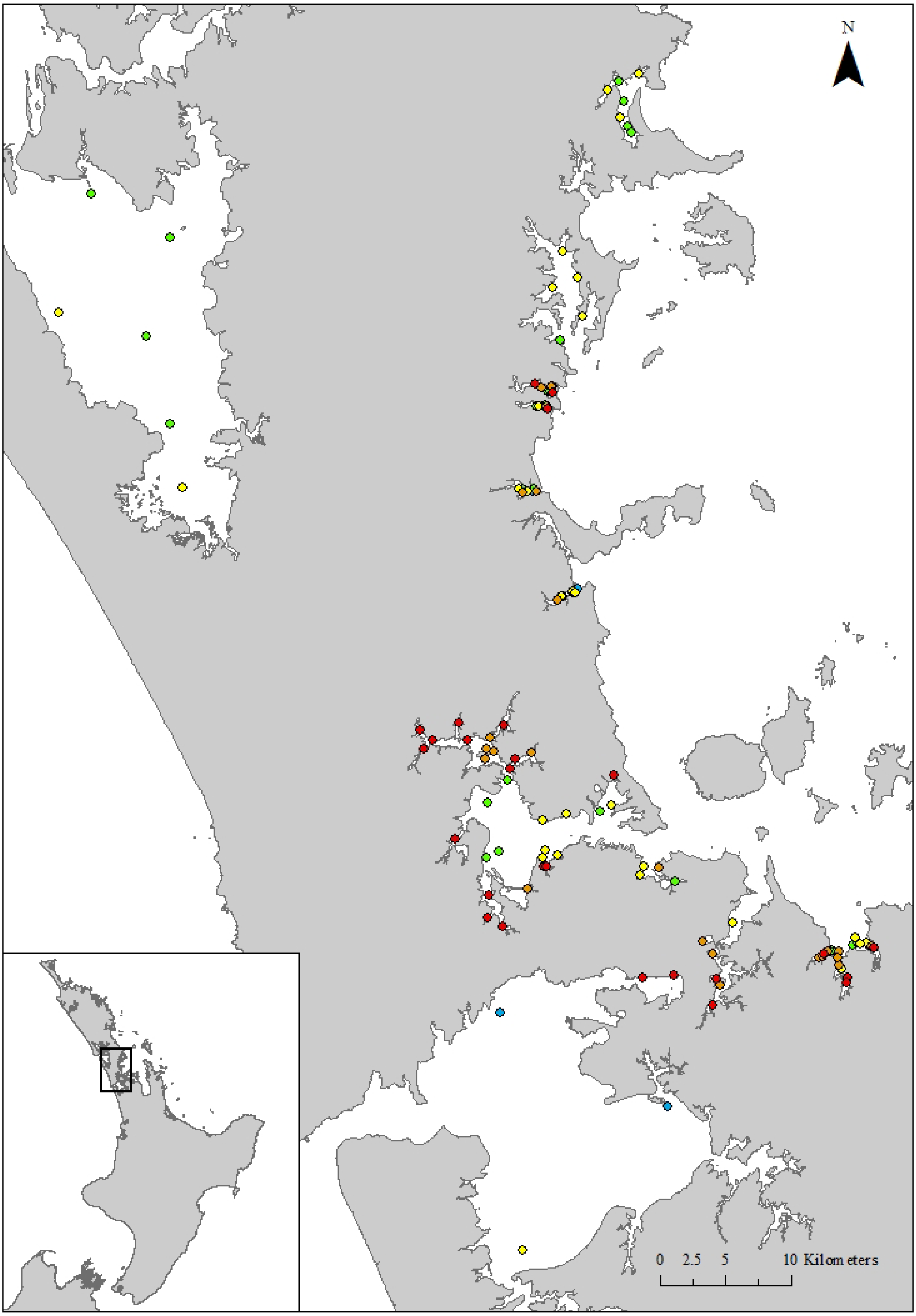


Figure 5-5 Upper Waitemata Harbour report card area, with the health scores of individual sites indicated by colour (red: unhealthy).

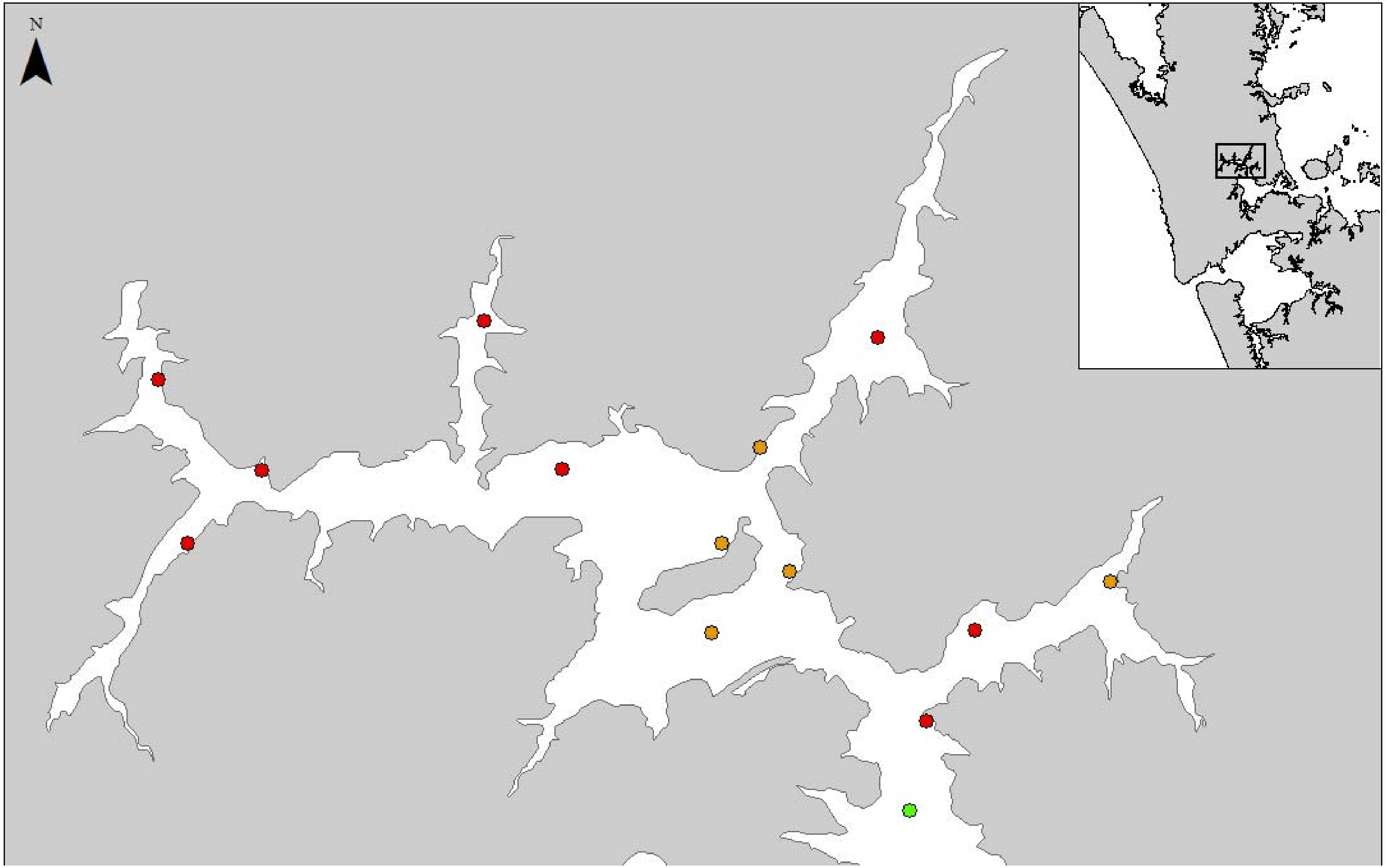


Figure 5-6 Central Waitemata Harbour report card area, with the health scores of individual sites indicated by colour (red: unhealthy).

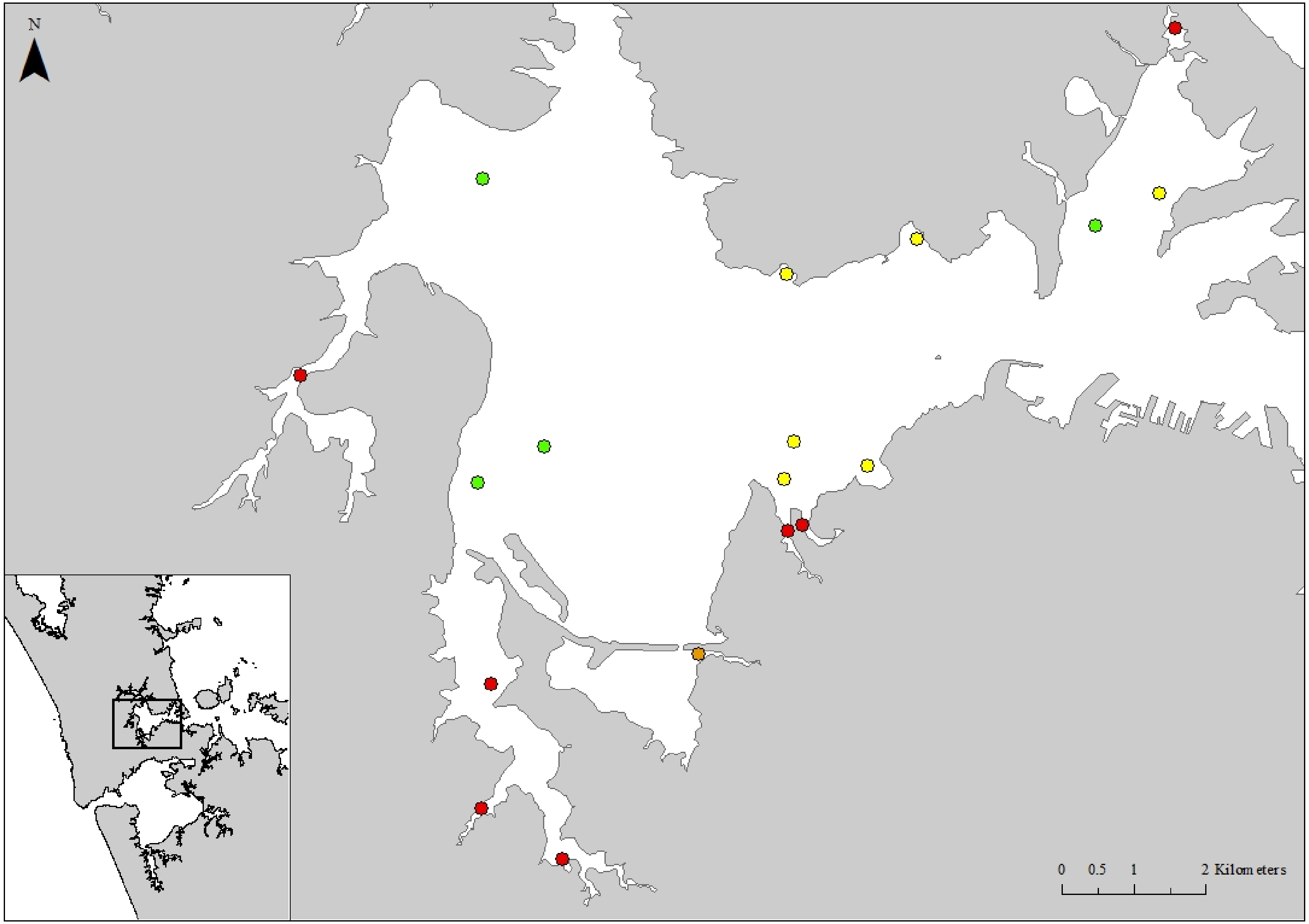
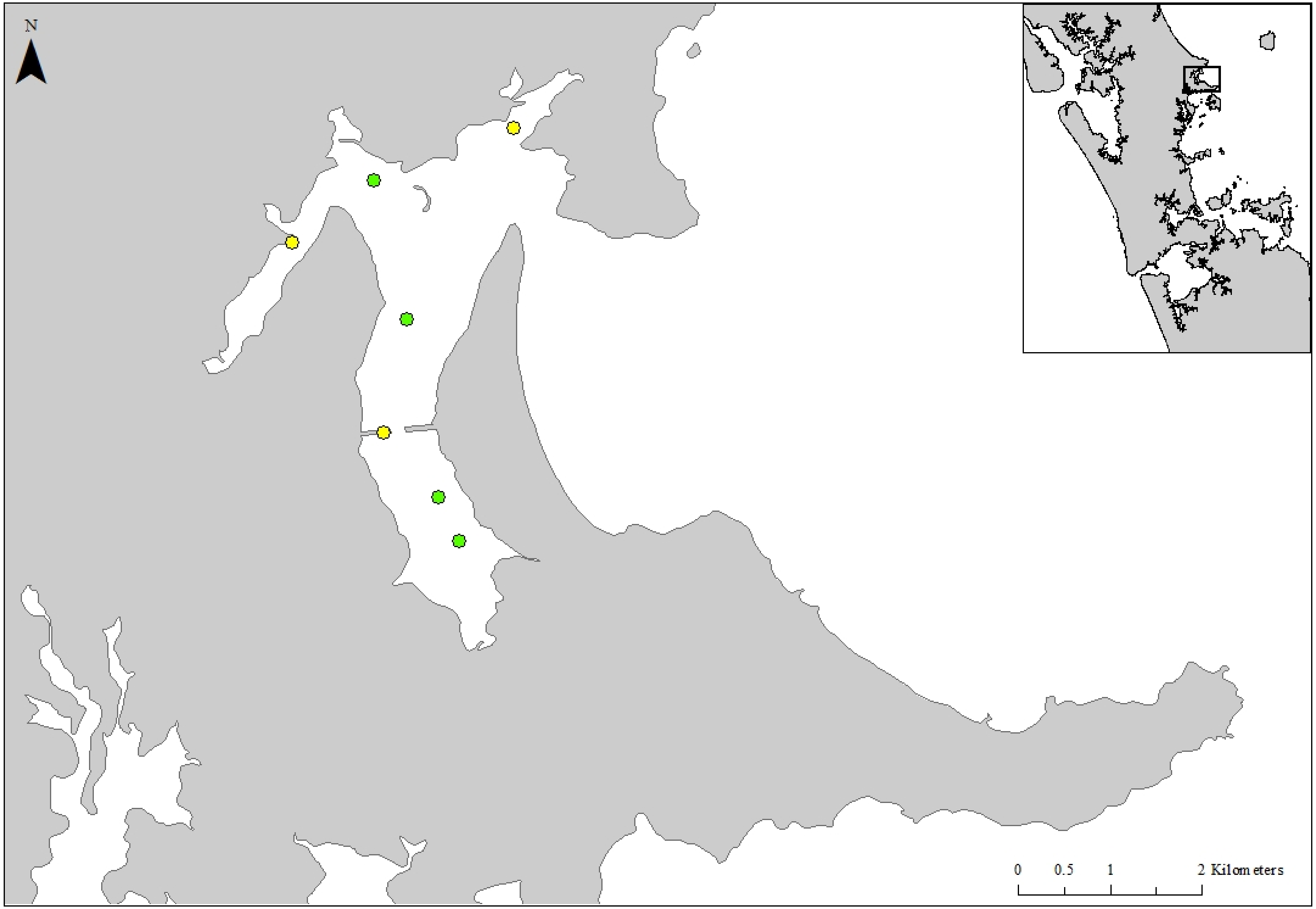


Figure 5-7 Whangateau Estuary, and the only estuarine sites within the Wellsford/Warkworth report card area. Health scores are indicated by colour (green: good).



6.0 Potentially Anomalous Scores

6.1 TBI

To identify sites with potentially anomalous TBI scores, we rank ordered the scores and plotted them alongside the names of the sites, their CAPmetals group numbers, and their colour codes (e.g., Figure 6.1). Three sites that are now sampled as part of the Estuarine Monitoring programme had anomalously low TBI scores (Table 6.1). Turanga J was lowest overall at 0.11; Waiwera J was 5th lowest at 0.17; Puhoi H was 11th lowest at 0.25, despite all three sites having CAPmetals and CAPmud scores that were ranked into groups 3 or 2. The low TBI scores at these sites are possibly a reflection of low salinity; but most likely be poorly resolved taxonomic identifications (these sites were among the first sampled in the RDP programme). For Turanga J, it was probably also related to the high sediment mud content (72.7%).

Table 6-1 Three sites with anomalously low TBI scores.

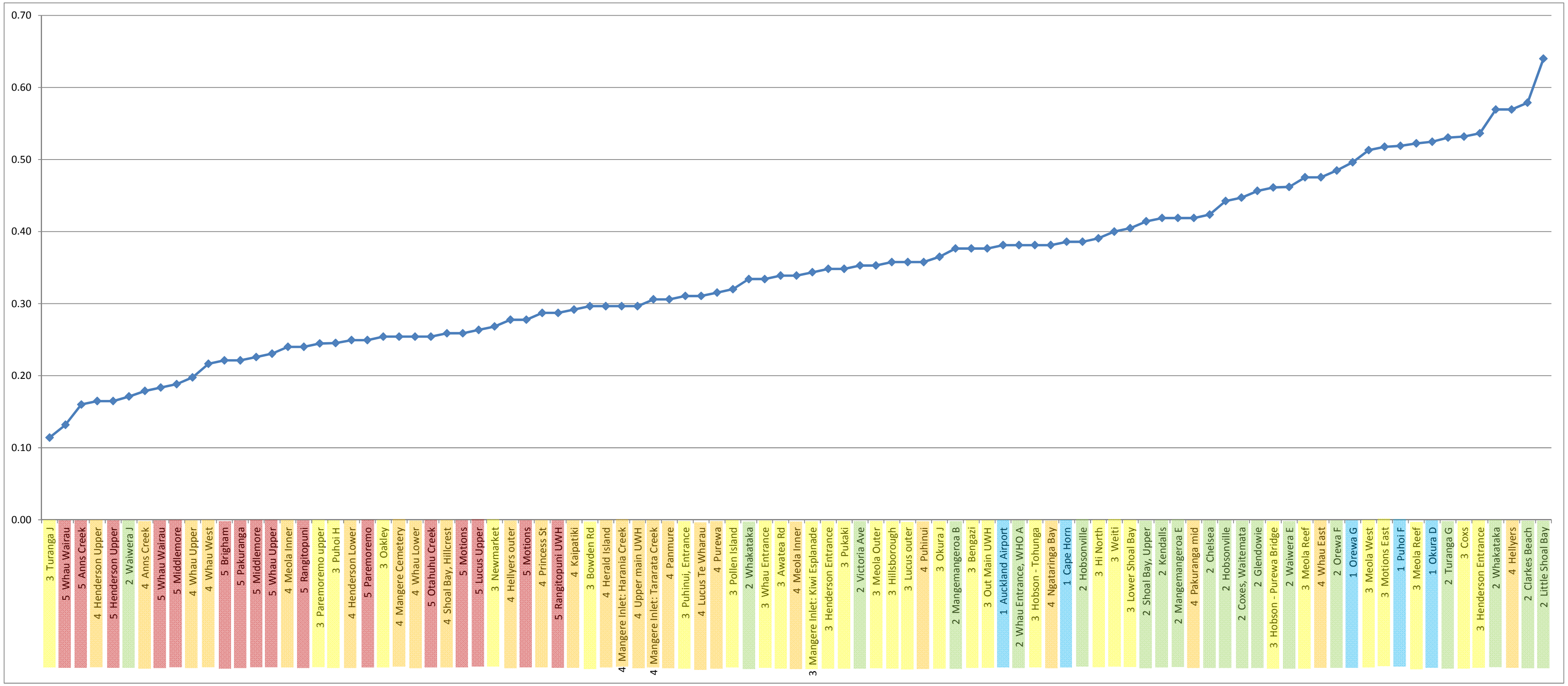
Site	PC1.500	MUD%	TBI	CAPmetal	CAPmud	GROUP
Turanga J	-0.051	72.7	0.11	-0.014	-0.010	3
Waiwera J	-1.976	14.6	0.17	-0.089	0.0265	2
Puhoi H	-1.499	22.0	0.25	-0.0344	0.0062	3

Pakuranga Mid, Whau East and Hellyers are all sites with relatively high heavy metal and mud concentrations that nevertheless had very high TBI scores (Table 6.2). Based on the mud and metals concentrations at these sites, the BHM scores appear to be accurate (although it should be noted that they are only marginally over the Group 3-4 cut-off). Thus, whilst the TBI scores may seem anomalously high, they may in fact provide useful information about remaining levels of functional redundancy at these sites. To understand factors responsible for the high TBI scores, it is necessary to examine the raw macrofaunal community data. The types of taxa that will make the largest contributions to TBI scores are moderately sized, worm-shaped animals living in fixed tube structures extending through the upper 0-2 cm of the sediment column. Examples include phoronids and polychaetes from the families Terebellidae, Trichobranchidae, Pectinariidae, Ampharaetidae, Maldanidae, Sabellidae, and Oweniidae. The spionid polychaete Boccardia syrtis forms high-density mats that can persist in relatively high mud content. These mats can attract other tubiculous species, which may have inflated the TBI scores in muddy areas that would otherwise score poorly (though we did not interrogate the raw community data to investigate the causes of anomalies at the three sites in Table 6.2). A recent invader is the spionid, Polydora cornuta, which also forms dense mats and appears more resistant to both mud and contaminants.

Table 6-2 Three sites with anomalously high TBI scores.

Site	PC1.500	MUD%	TBI	CAPmetal	CAPmud	GROUP
Pakuranga Mid	0.970	34.56	0.42	0.047	0.0613	4
Whau East	1.575	41.48	0.48	0.044	0.0609	4
Hellyers	0.463	53.43	0.57	0.044	0.0583	4

Figure 6-1 TBI scores at each of the original RDP sites sampled between 2002 and 2005. The CAPmetals group numbers are written before each site name, and the sites are colour coded accordingly. Note the predominance of red sites to the left and blue sites to the right, but the lack of perfect concordance.



RDP site names, with CAPmetals Group numbers and colour codings

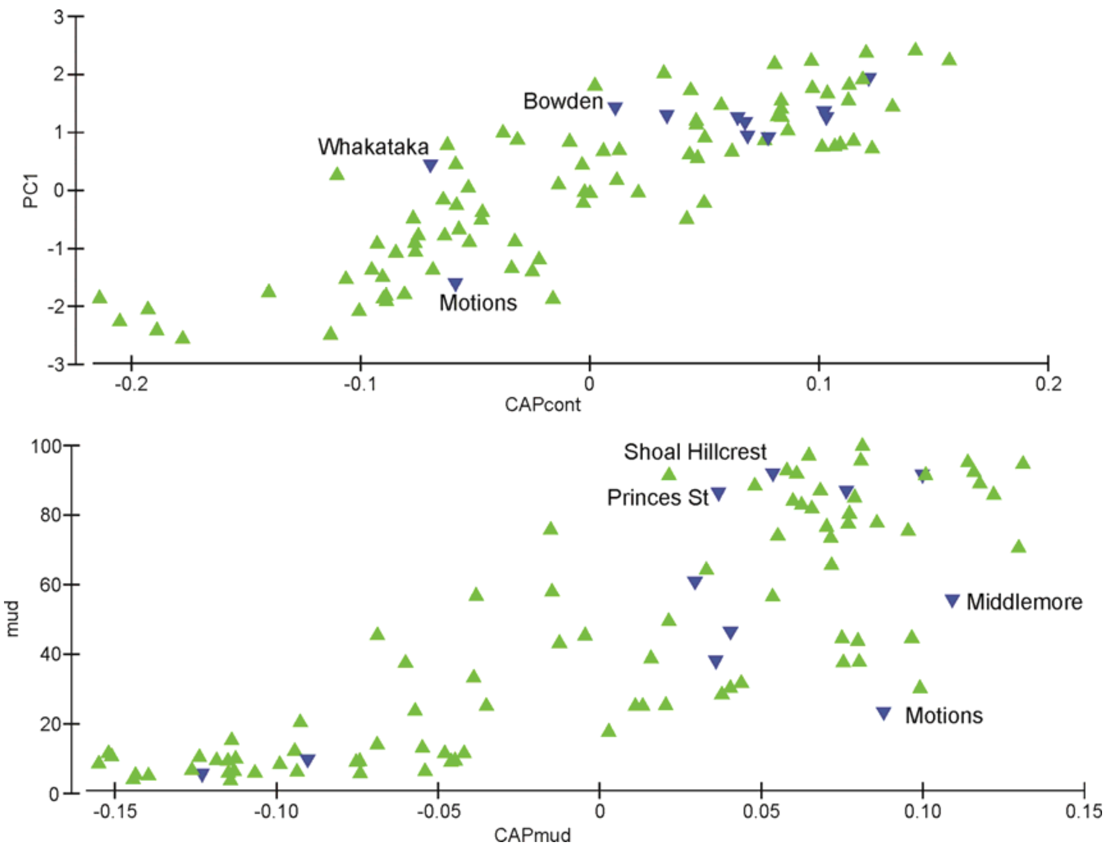
6.2 BHM

When assessing trends, we noticed a number of sites with much higher (or lower) CAPmetals and CAPmud scores for the first sampled date (Table 6.3). Generally, both CAPmetals and CAPmud scores exhibited the same pattern, but not always. The model CAP scores were plotted against the PCaxis of metal contamination or the mud content of the sediment to determine whether these point fell towards the outer edge of the data cloud, i.e., were outliers (Figure 6.2). This was not generally the case, although for CAPmetals, Middlemore, Motions and Newmarket were towards the edge. For CAPmud, Motions was again near the edge of the data cloud, as were Whakataka and Bowden.

Table 6-3 Sites whose initial sampling showed anomalously high or low CAPmetals or CAPmud scores.

Site	Year	CAPmetals	CAPmud
Bowden	2004		high
Shoal Hillcrest	2004		low
Mangere Cemetery	2005	Low	low
Middlemore	2002	Low	low
Motions	2002	low	
Newmarket	2005	High	high
Otahuhu	2004	Low	low
Panmure	2002		high
Princes St	2004	Low	low
Purewa	2004	Low	low
Whakataka	2002	low	low
Meola Outer	2004	high	
Lucas Outer	2005	low	

Figure 6-2 Position of initial sample points (dark blue) with anomalously high or low CAPmetals or CAPmud.



Some other anomalies were noted: the last value (2011) for CAPmud at Herald Island Waiarohiawas high, and CAPmetals values from MainU and MainO were very variable and had low values for 2006 and 2007 respectively.

All these sites exhibiting anomalies in their CAPmetals and CAPmud scores over time were compared with trends up to 2010 detected by the contaminant monitoring programme. Trends in contaminants were detected for Mangere Cemetery, Motions, MainO and Purewa, but not for CAPmetals and it is possible that these anomalies affected the ability of the trend analysis run on the CAPmetals scores to detect changes. However, it is also possible that these differences are due to the generally lower frequency of biological sampling affecting the power of the statistical analyses. With more data over time, it will be possible to correct for both these possibilities.

Generally, the results from the SOE monitoring programmes fit within the BHM framework, even when the sites sampled come from well outside the original spatial extent of sites in the model (e.g., Southern Kaipara sites). However, some of the sites from the estuarine monitoring programme were allocated higher CAPmetals scores than would have been anticipated from their heavy metal concentrations (Whangateau, Waikopua, Orewa and Turanga). This is most likely due to the lower degree of replication in this monitoring programme (6 replicates vs 10) affecting the overall number of taxa. However, it may also be due to slight differences in the sediment composition, with lower mud content. If it is the latter, the finding fits well with the results in Section 4 on the TBI, where sites with very low mud content had lower TBI scores, associated with lower species richness.

7.0 Comparisons With Contaminants

Comparisons with the trends observed by Mills et al. (2013) were made for all sites at which sufficient data were available for trend analysis (Table 7.1). In no cases were trends in CAPmetal scores observed to be contrary (in the opposite direction) to trends observed in total heavy metals. At two Upper Waitemata Harbour sites, Brighams and MainO (Outer main channel), both positive and negative trends in heavy metals were observed: no trend in CAPmetal scores was observed at MainO but a trend in the same direction as that of the strongest metal trend was observed at Brighams. Generally, more trends were detected in heavy metals (11 -14 statistically significant sites for the three metals, from Tables 6.4 to 6.6 in Mills et al. 2012) compared to trends in CAPmet. However, most of these were for the SOE sites, where the length of the time series differed for the two datasets. There are four reasons why this could occur:

1. Community composition response is more of a threshold response, thus recovery requires metal concentrations to drop below a certain level before it will recover. Similarly, with increasing contamination, in non-pristine areas, the community will already have altered to one adapted to stress, and thus will be able to cope with more stress for at least some while.
2. The time periods analysed are not the same, even for the Upper Waitemata Harbour sites, as the analyses considered in this report include the 2011 year. Some temporal patterns were not considered trends as the final sampling points indicated a cycle. Trend determination over short time periods always has the potential for trends to be identified that are really part of longer-term cycles (e.g., Stewart et al. 2013).
3. Outliers on the first sampling occasion which prevented detection of any trend as mentioned in section 7.
4. Different techniques were used for detecting trends in contaminants and CAPmetal scores and it is possible that the technique used for contaminants was more powerful. In future, both techniques should be used to analyse both datasets.

Table 7-1 Estimates of change over the monitored period (Relative Sen Slope Estimates as % of median per year) in statistically significant heavy metal contaminant concentration trends observed by Mills et al. (2013) and trends in CAPmetals scores from the BHMmetals. The sites presented in the table are those where sufficient data were available for trend analysis of CAPmetals.

		Copper	Lead	Zinc	CAPmet
Bowden	RDP			3.2	
Coxs	RDP			3.7	
Panmure	RDP			3.7	0.006
Kendall	RDP	1.8			
Whau Entrance	RDP		3.4	5.1	

Chelsea	RDP	6.4	3.5	2.5	
Princes	RDP				
Shoal Hillcrest	RDP				
Purewa	RDP				
Benghazi	RDP				
Otahuhu	RDP				
Whau Lower	SoE		-3.9		-0.01
Hobson Newmarket	SoE		-3.5		-0.025
Middlemore	SoE			1.7	
Motions	SoE	-4.2	-5.7		
Whau Wairau	SoE	-2.9	-4.3		
Oakley	SoE	-2.7	-2.4		
Henderson Upper	SoE	-3.9		-1.6	
Anns	SoE	-8.4	-6	-3.7	
Mangere Cemetery	SoE	-5.5	-4.5	-2.9	

		Copper	Lead	Zinc	CAPmet
Meola Inner	SoE	-2.7	-3.4	-1.5	
Whau Upper	SoE		-2.9	1.2	
Meola Reef Te Tokaroa	SoE				
Rangitopuni UWH	UWH			3	
Herald Island North	UWH	-14.2			
Central Main Channel	UWH	-4		2.7	
Brighams UWH	UWH	-3.1		2.9	-0.013
Hellyers Upper UWH	UWH				0.015
Herald Island Waiarohia	UWH				
Lucas UWH	UWH				-0.015
Outer Main Channel	UWH				
Upper Main Channel	UWH				

8.0 Conclusions and Recommendations

The results of this report confirm that the original BHMmetals, the recently developed BHM mud and TBI are all still working to give consistent results. The indices are robust with respect to long-term cycles caused by climate patterns, yet some trends have been detected.

The larger degree of temporal variability that occurs over time for the CAPmud scores compared with the CAPmetals scores raises an interesting possibility. Changes over time in macrofaunal assemblages have been observed in Manukau related to ENSO and changes in exposure to wind waves (Hewitt and Thrush 2009, Turner et al. 1995). Obvious differences between seasons and between years can also be seen in water column visibility. Therefore, it is possible that the macrofaunal communities are responding to small changes in mud that are deposited and then eroded, or to changes in suspended sediment concentrations. If this is the case, this may give us a new direction of research that would address a previously highlighted question: how can we determine changes related to terrestrial sediment in areas where erosion of sediments rather than deposition is the predominant process?

Particularly important for the success of the BHM model in associating changes in community composition to either contaminants or mud content is the finding that only rarely do we observe changes in both CAPmetals and CAPmud scores simultaneously. Occasionally this does happen (e.g., at Brighams), but, given that contaminants can be attached to sediment it is not surprising that occasionally contaminants and sediment mud content could be changing simultaneously.

The BHMmetals is based on a PCA axis of total copper, lead and zinc, as this technique deals with the correlation between the metal concentrations. At the time, it was known that the lead concentrations in the seafloor sediments were likely to decrease now that lead was no longer used in petrol. Mills and Williamson (2012) document decreases in lead concentrations at a number of sites, however we found no evidence that this change was affecting the ability of the BHMmetals. We recommend that future reports continue to analyse for this potential problem, but as no effects were observed, we do not consider it is necessary to resample these sites. Moreover, we also found relatively little change as yet in sites where mangrove encroachment is occurring. Thus, there is no need for a special sampling trip to relocate these sites. However, over the longer-term it would be appropriate to relocate these sites to areas lacking mangroves. When this is done, concurrent sampling at both the old and new site would be useful to allow any previous changes to be tracked though to the new site.

In the compilation of this report, it became apparent that having the BHM scores run from -0.25 (very good) to +0.25 (unhealthy), while the TBI ran from 1 (good) to 0 (poor) was confusing. Even on its own the BHM was awkward, as negative trends are reported as improving health, although this does fit well when comparing with heavy metal trends (i.e., decreasing trends in heavy metal concentrations are comparable to decreasing trends in CAPmetal scores). For this reason we recommend that in future (2014 reporting onwards) the BHM scores are standardized to run between 5 (very healthy) and 0 (unhealthy) by adding 0.25 and multiplying by -10. This score should be reported as CAP standardised scores to clearly differentiate them from the previous values.

The BHMmetals, BHM mud and TBI all produce complementary information (composition, functionality and resilience), with sensitivity ranges that are also complementary. The TBI also integrates across the interactions between the effect of heavy metals and mud on the macrofaunal communities. For these reasons, we recommend using all three indices. We suggest that average health values can be determined for a site in the following way:

1. If the CAPmud score allocated the site to Mud group 1 then Health is calculated as the average CAPmetals and CAPmud group scores.
2. If the CAPmetals score allocated the site to Cont group 4 or 5 then Health is equal to the TBI group score.
3. Otherwise, Health is the average of the CAPmetals, CAPmud and TBI group scores.
4. Recoding these scores as:
    a. $\leq$ 0.2 " extremely good"
    b. 0.2 – 0.4 inclusive "good"
    c. 0.4 – 0.6 exclusive "moderate"
    d. 0.6 – 0.8 exclusive "poor"
    e. $\geq$ 0.8 "unhealthy with low resilience"

Amalgamating health scores across a whole report card area is considered unwise, as not all muddy, healthy or contaminated areas have been sampled proportionally. However, the GIS plots we drew of the individual sites and their rankings manage to show how much has been sampled, approximate sample locations and still give information on the overall health. Amalgamating changes over time is less problematic. A report card could simply show a pie chart with segments of red (proportion getting worse), yellow (staying the same), green (getting better). The yellow segment should probably be split into a blue segment (already ranked as good) and an orange segment (staying the same and not good).

The GIS plots of the report card areas clearly demonstrate the uneven coverage of sampling sites within the different areas (e.g., Figure 5.4). Admittedly, the focus is on sampling in areas where problems from storm-water contaminants are expected. Other sites are located in areas where State of the Environment monitoring has been established for other reasons, e.g., Manukau, Mahurangi and Kaipara harbours and the major estuaries along the East Coast. However, there are some obvious gaps, e.g., Bethels Beach estuary, Muriwai, Piha, Wairoa, Kawakawa Bay, Huruhi Bay and Okahuiti Creek on Waiheke and Matakana estuary.

Comparison of trends in community composition and stormwater-contaminants were generally consistent with one another, although more trends were detected for stormwater contaminants than for community composition. The differences in the number of trends detected may reflect both biological responses (i.e., namely that biological responses may lag behind changes in contaminant chemistry and there may be hysteresis) as well as differing trend detection methodologies, which we recommend to be standardized in the future. We also suggest that trends in the TBI which occur above 0.4 and below 0.3 may be considered uninformative, and only trends approaching or crossing these boundaries are important.

9.0 Acknowledgements

We thank Scott Edhouse, Katie Cartner and Julia Simpson for help with figures and GIS. Many thanks to the AC for funding this work and the collection of the original data sets.

10.0 References

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

Anderson, M. J., Willis, T. J. 2003. Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology. *Ecology,* 84: 511-525.

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 NIWA for Auckland Regional Council. *Auckland Regional Council Technical Publication Number,* 317.

Anderson, M. J., Gorley, R. N., Clarke, K. R. 2008. *PERMANOVA for PRIMER: guide to software and statistical methods.* PRIMER–E Ltd, Plymouth.

Clarke, R. T., Gorley, R. N. 2006. Primer v6. PrimerE, *Plymouth Council Technical Publication,* 184.

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., Simpson, J. 2012. Assessment of the Estuarine Ecological Monitoring Programme to 2012. Prepared by NIWA for Auckland Regional Council. *Auckland Council Technical Report,* 2012/040.

Lohrer, D., Rodil, I. F. 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*.*

Lohrer, A. M., Hailes, S., Cartner, K., Wadhwa, S., Edhouse, S. 2012. Wairoa Embayment: Benthic habitats, environmental status and potential future threats. Prepared by NIWA for Auckland Council. *Auckland Council Technical Report,* 2012/

Mills, G., Williamson, B. 2011. Marine sediment contaminants: Status and Trends Assessment 1998 to 2010. Prepared by Diffuse Sources Ltd for Auckland Regional Council. *Auckland Council Technical Report,* 2012/041.

Stewart, M., Gadd, J., Ballantine, D., Olsen, G. 2013. Shellfish Contaminant Monitoring Programme: Status and Trends Analysis 1987 - 2011*. Prepared by NIWA for Auckland Council,* HAM2012-115: 208.

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.

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*

11.0 Appendix 1: Summary of health status across the report card areas

Year sampled	Site Name	NZTM East	NZTM North	Reporting Area	CAPmetal	CAPmud	TBI	Overall Health Score
2011	Hbv	1749660	5926258	Central Waitemata	-0.156	-0.148	0.56	0.30
2011	Whau	1748974	5920790	Central Waitemata	-0.117	-0.131	0.51	0.30
2010	Purewa	1762364	5918498	Central Waitemata	0.028	0.054	0.46	0.33
2011	ShB	1756642	5923858	Central Waitemata	-0.101	-0.073	0.69	0.38
2010	Hend_entr	1748127	5924512	Central Waitemata	-0.071	-0.110	0.54	0.38
2010	Wh_ent	1748063	5920287	Central Waitemata	-0.069	-0.102	0.48	0.38
2010	Coxes	1753479	5920531	Central Waitemata	-0.061	-0.092	0.41	0.44
2010	Kendall	1752352	5923186	Central Waitemata	-0.067	-0.017	0.53	0.44
2010	MOuter	1752317	5920334	Central Waitemata	-0.062	-0.117	0.45	0.44
2011	LoS	1757533	5924310	Central Waitemata	-0.045	0.011	0.46	0.51
2011	Awatea	1760037	5919688	Central Waitemata	-0.037	0.000	0.46	0.51
2010	Chelsea	1754161	5923677	Central Waitemata	-0.052	-0.014	0.47	0.51
2011	MReef	1752452	5920868	Central Waitemata	-0.049	0.007	0.46	0.51
2011	Newmarket	1759726	5918973	Central Waitemata	-0.044	-0.053	0.37	0.56
2011	Whakataka	1761184	5919536	Central Waitemata	-0.033	-0.001	0.39	0.62
2011	Oakley	1751121	5917912	Central Waitemata	0.027	0.042	0.3	0.67
2010	Hillcrest	1757749	5926606	Central Waitemata	0.067	0.061	0.28	1.00
2011	Hend_upp	1745597	5921791	Central Waitemata	0.052	0.066	0.2	1.00
2011	MInner	1752369	5919629	Central Waitemata	0.038	0.069	0.25	1.00
2011	Motions	1752573	5919704	Central Waitemata	0.055	0.053	0.26	1.00
2011	WhLower	1748243	5917496	Central Waitemata	0.036	0.046	0.26	1.00
2011	WhUpper	1749226	5915064	Central Waitemata	0.031	0.066	0.2	1.00

2011	WhWairau	1748106	5915757	Central Waitemata	0.105	0.067	0.21	1.00
2011	Okura1	1755005	5940844	Hibiscus Coast	-0.176	-0.129	0.37	0.20
2011	Okura4	1754576	5940526	Hibiscus Coast	-0.199	-0.136	0.46	0.20
2011	Orewa3	1751499	5948270	Hibiscus Coast	-0.203	-0.134	0.48	0.20
2011	Waiwera6	1752406	5954714	Hibiscus Coast	-0.185	-0.130	0.35	0.20
2011	Turanga1	1775931	5913657	Hibiscus Coast	-0.145	-0.143	0.32	0.30
2011	Okura7	1753871	5940246	Hibiscus Coast	-0.182	-0.110	0.62	0.31
2011	Orewa2	1751623	5948375	Hibiscus Coast	-0.126	-0.077	0.50	0.38
2011	Orewa5	1750795	5948424	Hibiscus Coast	-0.119	-0.054	0.45	0.38
2011	Waiwera2	1752718	5954601	Hibiscus Coast	-0.141	-0.065	0.56	0.38
2011	Waiwera5	1752522	5954834	Hibiscus Coast	-0.120	-0.050	0.47	0.38
2011	Waiwera9	1751853	5954750	Hibiscus Coast	-0.091	-0.073	0.43	0.38
2011	Okura3	1754577	5940596	Hibiscus Coast	-0.185	-0.118	0.35	0.42
2011	Waiwera8	1751980	5954664	Hibiscus Coast	-0.166	-0.106	0.38	0.42
2011	Okura2	1754767	5940513	Hibiscus Coast	-0.081	-0.013	0.52	0.44
2011	Orewa8	1750537	5948449	Hibiscus Coast	-0.110	-0.043	0.40	0.44
2011	Orewa4	1751194	5948184	Hibiscus Coast	-0.149	-0.082	0.34	0.49
2011	Waiwera4	1752419	5954658	Hibiscus Coast	-0.163	-0.087	0.32	0.49
2011	Okura8	1753697	5940183	Hibiscus Coast	-0.103	-0.038	0.36	0.56
2011	Turanga7	1775097	5911868	Hibiscus Coast	-0.078	-0.019	0.36	0.56
2011	Turanga3	1774900	5913136	Hibiscus Coast	-0.080	-0.114	0.26	0.60
2011	Waiwera3	1752622	5954701	Hibiscus Coast	-0.145	-0.083	0.23	0.60
2011	Orewa6	1750791	5948141	Hibiscus Coast	-0.030	0.010	0.37	0.62
2011	Turanga6	1774902	5912091	Hibiscus Coast	-0.062	-0.004	0.30	0.62
2011	Orewa1	1751782	5948250	Hibiscus Coast	-0.091	-0.025	0.22	0.67

2011	Okura9	1753490	5939953	Hibiscus Coast	-0.034	0.022	0.38	0.69
2011	Turanga4	1774798	5912683	Hibiscus Coast	-0.021	0.013	0.27	0.73
2011	Turanga8	1775506	5911190	Hibiscus Coast	0.007	0.061	0.29	0.80
2011	Waiwera1	1752683	5954504	Hibiscus Coast	-0.017	0.034	0.29	0.80
2011	Turanga10	1775466	5910834	Hibiscus Coast	0.035	0.093	0.19	1.00

2011	NPC	1723897	5953342	Kaipara	-0.188	-0.068	0.63	0.31
2011	TPB	1717930	5970907	Kaipara	-0.175	-0.066	0.74	0.31
2011	KaiF	1722114	5960073	Kaipara	-0.105	-0.062	0.52	0.38
2011	KKF	1723897	5967569	Kaipara	-0.127	-0.087	0.87	0.38
2011	HCK	1715422	5961841	Kaipara	-0.174	-0.066	0.31	0.42
2011	KaiB	1724876	5948515	Kaipara	-0.109	-0.048	0.53	0.44

2011	JB	1753681	5959768	Mahurangi	-0.099	-0.070	0.83	0.38
2011	DC	1753106	5963751	Mahurangi	-0.070	-0.065	0.39	0.49
2011	MH	1754969	5964505	Mahurangi	0.005	0.015	0.43	0.51
2011	TK	1755354	5961583	Mahurangi	-0.033	-0.003	0.48	0.51
2011	HL	1753798	5966477	Mahurangi	0.006	0.053	0.42	0.58

2010	AA139	1761854	5901338	Manukau	-0.226	-0.159		0.20
2009	CH133	1749082	5908472	Manukau	-0.222	-0.133		0.20
2011	ManAA	1761854	5901338	Manukau	-0.202	-0.150	0.38	0.20
2011	ManCB	1750772	5890380	Manukau	-0.063	-0.079	0.65	0.44
2011	Anns	1762281	5911361	Manukau	0.139	0.137	0.18	1.00
2011	Mangare	1759928	5911221	Manukau	0.098	0.109	0.19	1.00

2010	Benghazi	1766790	5915326	Tamaki Estuary	-0.004	0.013	0.4	0.51
2010	Bowden	1765251	5912952	Tamaki Estuary	0.073	0.061	0.36	0.67
2010	Panmure	1764477	5913898	Tamaki Estuary	0.063	0.056	0.34	0.67
2010	Princes	1765853	5910587	Tamaki Estuary	0.082	0.075	0.32	0.67
2010	Otahuhu	1765518	5911051	Tamaki Estuary	0.071	0.048	0.24	1.00
2011	Middlemore	1765216	5909093	Tamaki Estuary	0.077	0.058	0.2	1.00
2011	Waikopua3	1776332	5913863	Tamaki Strait	-0.143	-0.123	0.44	0.30
2011	Mangeman1	1774436	5913619	Tamaki Strait	-0.107	-0.068	0.62	0.38
2011	Mangeman2	1774297	5913309	Tamaki Strait	-0.138	-0.107	0.48	0.38
2011	Mangeman3	1774206	5913215	Tamaki Strait	-0.110	-0.067	0.61	0.38
2011	Waikopua1	1776112	5914191	Tamaki Strait	-0.114	-0.049	0.67	0.44
2011	Waikopua4	1776463	5913782	Tamaki Strait	-0.077	-0.075	0.34	0.49
2011	Waikopua6	1776964	5913829	Tamaki Strait	-0.093	-0.052	0.38	0.49
2011	Waikopua7	1777116	5913628	Tamaki Strait	-0.094	-0.059	0.34	0.49
2011	Mangeman5	1774061	5913223	Tamaki Strait	-0.014	0.054	0.40	0.58
2011	Mangeman6	1774018	5913184	Tamaki Strait	-0.043	0.014	0.37	0.62
2011	Waikopua8	1777378	5913554	Tamaki Strait	-0.080	-0.018	0.26	0.67
2011	Mangeman10	1773392	5912719	Tamaki Strait	0.025	0.090	0.30	0.67
2011	Mangeman4	1774120	5913202	Tamaki Strait	0.005	0.073	0.31	0.69
2011	Mangeman9	1773601	5912824	Tamaki Strait	-0.016	0.063	0.33	0.69
2011	Mangeman7	1773706	5912990	Tamaki Strait	-0.030	0.057	0.29	0.80
2011	Mangeman8	1773572	5912856	Tamaki Strait	-0.009	0.071	0.29	0.80
2011	Waikopua9	1777523	5913463	Tamaki Strait	0.026	0.051	0.23	1.00
2011	HIN11	1747994	5928628	Upper Waitemata	-0.035	-0.044	0.38	0.62

2011	HIW11	1747901	5927833	Upper Waitemata	-0.052	-0.047	0.36	0.62
2011	MainO11	1748592	5928382	Upper Waitemata	-0.028	-0.018	0.34	0.62
2009	HellU9	1751444	5928286	Upper Waitemata	0.094	0.062	0.31	0.67
2009	LucU9	1748335	5929477	Upper Waitemata	0.007	0.030	0.35	0.69
2011	BRIG11	1743254	5928631	Upper Waitemata	0.005	0.040	0.13	0.80
2011	MainC11	1746577	5929280	Upper Waitemata	0.011	0.044	0.26	0.80
2011	Paremoremo	1745881	5930603	Upper Waitemata	0.069	0.069	0.27	1.00
2011	Hell11	1750242	5927860	Upper Waitemata	0.039	0.030	0.27	1.00
2011	Luc11	1749374	5930448	Upper Waitemata	0.053	0.063	0.27	1.00
2011	MainU11	1743908	5929274	Upper Waitemata	0.025	0.059	0.22	1.00
2011	OHBV11	1749807	5927056	Upper Waitemata	0.033	0.038	0.20	1.00
2010	RNG10	1742993	5930083	Upper Waitemata	0.059	0.079	0.15	1.00
2011	Puhoi4	1752848	5955879	Wellsford/Warkworth	-0.154	-0.119	0.55	0.38
2011	Puhoi6	1752639	5955816	Wellsford/Warkworth	-0.161	-0.105	0.50	0.38
2011	Whangate1	1759055	5975559	Wellsford/Warkworth	-0.074	-0.070	0.44	0.38
2011	Whangate2	1758828	5976034	Wellsford/Warkworth	-0.126	-0.095	0.40	0.38
2011	Whangate4	1758492	5977967	Wellsford/Warkworth	-0.073	-0.103	0.49	0.38
2011	Whangate7	1758136	5979479	Wellsford/Warkworth	-0.081	-0.088	0.62	0.38
2011	Puhoi1	1753095	5956252	Wellsford/Warkworth	-0.170	-0.113	0.35	0.42
2011	Whangate3	1758234	5976733	Wellsford/Warkworth	-0.063	-0.059	0.72	0.44
2011	Whangate6	1759657	5980044	Wellsford/Warkworth	-0.065	-0.050	0.46	0.44
2011	Whangate5	1757249	5978798	Wellsford/Warkworth	-0.036	-0.044	0.42	0.51
2011	Puhoi7	1752210	5956173	Wellsford/Warkworth	-0.034	-0.005	0.36	0.62
2011	Puhoi3	1752977	5956246	Wellsford/Warkworth	-0.037	0.036	0.33	0.69
2011	Puhoi2	1753047	5955757	Wellsford/Warkworth	-0.011	0.040	0.23	0.80
2011	Puhoi9	1751705	5956419	Wellsford/Warkworth	-0.028	0.050	0.28	0.80