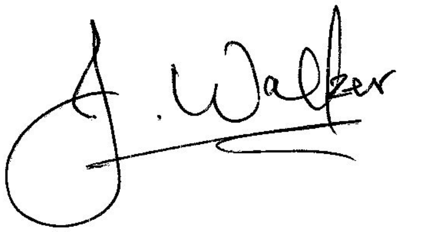
Auckland
Regional Council
TE RAUHĪTANGA TAIAO

Assessment of the Benthic Health Model

June 2010 TR2010/034

Auckland Regional Council
Technical Report No.034 August 2010
ISSN 1179-0504 (Print)
ISSN 1179-0512 (Online)
ISBN 978-1-877540-90-5

Reviewed by:



Name: Jarrod Walker

Position: Project Leader - Marine

Organisation: Auckland Regional Council

Date: 28-10-2010

Approved for ARC Publication by:

Name: Grant Barnes

Position: Group manager Monitoring and Research

Organisation: Auckland Regional Council

Date: 28-10-2010

Recommended Citation:

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

Assessment of the Benthic Health Model

Judi Hewitt
Joanne Ellis

Prepared for
Auckland Regional Council

NIWA Client Report: HAM2010-124
June 2010

NIWA Project: ARC 10218

National Institute of Water & Atmospheric Research Ltd
Gate 10, Silverdale Road, Hamilton
PO Box 11115, Hamilton, New Zealand
Phone +64-7-856 7026, Fax +64-7-856 0151
www.niwa.co.nz

Reviewed by:

Dr Drew Lohrer

Approved for release by:

Dr Simon Thrush

Formatting checked

A. Bartley

Contents

1 Executive Summary

The Benthic Health Model has been developed by the ARC to provide a tool for classifying intertidal sites within the region according to categories of relative ecosystem health, based on its community composition and predicted responses to storm-water contamination. The model is a multivariate analysis (CAP) of macrobenthic community composition backed by information on sediment copper, lead and zinc concentrations.

While the Benthic Health Model has proven to be a useful tool, three issues have been identified by the ARC that would facilitate their use of the model as a State of the Environment (SOE) indicator:

- The predicted response to storm-water contamination needs to be broadened to include other anthropogenic influences, specifically changes to sediment mud content.
- The effects of occasional incorrect site assignments in the BHM, specifically those driven by the recruitment of juveniles prior to sampling, need to be determined.
- The sensitivity of the model to detect changes associated with increased or deceased pollution needs to be ascertained, and the effects of adding new sites, times and species to the existing model need to be determined.

A new model was derived that associated strong gradients of community change in response to mud content of the sediment using the BHM data (referred to as CAPMud). Due to some level of correlation between effects related to CAPMud and those related to sediment stormwater contamination (referred to as CAPMetal), we recommend that bivariate plots of CAPMud and CAPMetal be used to assess site changes and hence species responses to the two varying anthropogenic stressors. For example movement of a site along the CAPMetal axis would suggest effects associated with contaminants. Movement along the CAPMud axis would indicate species responses to elevated mud content. Site changes in both directions may be a result of responses to both factors and would initiate a close inspection of which species are responding to the changes.

Differences in response times of different life stages of bivalves driven by dispersal were identified as a potential confounding influence on the BHM. At present, the problem generated by dispersal can be dealt with by investigating obvious outliers in for high abundances of bivalves and then determining whether the majority of these were juveniles. However, if size class information for bivalve species was routinely recorded, a new BHM could be developed in the future that was less susceptible to dispersal of juveniles.

The BHM could show simulated changes in community composition as low as 5% for increased health and as low as 10% for decreased health. No movement between rankings was observed, with the model providing a more sensitive assessment of change if actual movement along the CAP axis is calculated.

Adding new sites into the model had little effect on the strength of the relationship between community composition and contamination, suggesting that the model could be updated without weakening it. However, a number of sites exhibited changes of greater than 5% in their CAP score, suggesting that new sites should be assessed one at a time rather than in groups. Conversely, for existing model sites, new times can be assessed simultaneously.

There was relatively little effect on the BHM of the addition of new species into the model, though the new species that were available for testing this effect were not highly abundant. In cases where new species are highly abundant or invasive, the strength of this effect on the BHM may need to be reassessed.

2 Introduction

In 2001, the Auckland Regional Council (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. As a result of the workshop, research initiatives to develop relevant criteria of health for urban marine coastal areas were commissioned. The research utilised existing data of estuarine benthic communities and associated impacts from storm water inputs and resulted in the development of a multivariate model of ecosystem health (Anderson et al., 2002, ARC 2002, Hewitt et al., 2005). The purpose of the model was to provide a tool whereby new observational data of the community at a given site within the region could be classified, using only this biological information, into a category of relative ecosystem health. These models of "health" relied on the definition of the rank pollution of sites along a gradient (where 1 = "healthy" and 5 = "polluted"), which was defined indirectly through the analysis of sediment chemistry and existing knowledge by ARC managers. Therefore it was proposed that the benthic health model be based directly on quantitative information regarding the chemistry at particular sites. Also, sites that form the basis for models should be of varying sediment characteristics and varying concentrations of contaminants in intertidal estuarine environments that are interspersed and that span the region of interest. Thus, appropriate existing data were assembled and more data were collected, which included both biological information and the chemical and physical characteristics of sediments, in order to develop new models (Anderson et al., 2006). The resultant Benthic Health Model (BHM) assesses health of a site based on its community composition and predicted responses to storm-water contamination. Ecological assemblages were found to reflect pollution gradients very well, and the study was able to identify clear methods for modeling pollution gradients using ecological data (Anderson et al., 2006).

Reports on the status of near-shore marine sediments at sites representing stormwater receiving environments have been prepared in 2003, 2005 and 2009 using the Benthic Health Model developed by Anderson et al., (2006). In Tricklebank 2010, links between contamination gradients, sediment particle size and possible descriptors of ecological assemblages including biodiversity indices and individual taxa were examined. Again there was a strong match between the levels of contaminants found in sediments and the 'health' of the community of marine organisms that resided in sediments at each site (Tricklebank et al., 2007).

At present the Benthic Health Model has proven to be a useful tool for assessing the health of a site based on its community composition and predicted responses to storm-water contamination. However based on the existing research three issues were identified that would aid in the refinement of the BHM:

- In order to make the model more useful as a State of the Environment (SOE) indicator the predicted response to storm-water contamination needs to be broadened to include other anthropogenic influences. Therefore in this report we investigate the inclusion of sediment data (which can reflect terrigenous sediment inputs) into the BHM.

- There have been occasional incorrect site assignments in the BHM driven by recruitment of juveniles occurring prior to sampling. Therefore we wished to determine whether separating these species into different size classes and keeping the juveniles separate for the community analysis could prevent future problems.
- Sampling has now been conducted over seven years and hence a review of the model is required. In this report we review the BHM and consider; i) model sensitivity to detect changes in sites' health status over time, which is required if the ARC is to use the BHM as a diagnostic and monitoring tool, and ii) the effect of introducing new species, sites and times by creating a series of new models of varying degrees of change.

Below we discuss the methods and results for each of these three sections in turn.

3 Using the Benthic Health Model to determine responses to changes in sediment mud content

3.1 Methods

The BHM is a multivariate model of community structure, based on canonical analysis of principal coordinates (CAP). The full model development is outlined in Anderson et al., (2002). Data assembled from sites across the Auckland Region included mean abundances of 103 taxa from 84 sites, some of which were sampled in multiple years (from 2002-2005), yielding 95 samples. The original models developed by Anderson et al., (2002) used 81 samples (M), with 14 samples being reserved to provide independent model validation (V). In this report we develop models using the 81 model sites (M). Chemical data consisted of measures of concentrations (mg/kg) of copper, lead and zinc from the total sediment sample (< 500 m) and also from weak acid extraction of the mud fraction (< 63 m). Anderson et al., (2006) determined that the biotic assemblages had the strongest relationships with metal concentrations in the total sediment sample (PC1.500), rather than in the mud fraction (PC1.63). Further Anderson et al., (2006) compare various dissimilarity matrices and subsets of taxa in terms of the strength of their relationship with the pollution gradients. Bray-Curtis for square-root transformed abundances achieved the best relationship with the PC1.500 pollution gradient (Anderson et al., 2006).

While the primary sediment characteristic of interest to the ARC was sediment mud content, initially we determined which sediment grain size characteristics were likely to be the most important for macrofaunal community composition. Correlations between community composition and sediment grain size patterns were examined using PRIMER's BIO-ENV, which uses all the available environmental variables to find the combination that 'best explains' the patterns in the biological data. A Spearman rank correlation between the species data, using Bray-Curtis similarities, and the sediment grain size (Euclidean distance for four grain size variables; coarse sand, medium sand, fine sand and mud fraction) was performed.

Following this, sediment mud content was incorporated into the existing model by producing a separate CAP axis. This analysis was conducted on the data used to create the original model (81 M samples) using Bray-Curtis dissimilarity of square root transformed data and the mud content from sediment grain size data. The multivariate computer software package, PRIMER v6, with the add-on PERMANOVA+ (Anderson and Gorley, 2007) was used to analyse the data.

The potential for interactions between the separately derived CAP axes (i.e., mud versus contaminants) was explored in three ways. (1) A Pearson's correlation was run between the CAP scores along the Mud and the Metal axes. (2) The BIO-ENV analysis was run a second time using not only the sediment grain size characteristics but also PC1.500. (3) Partial canonical correspondence analysis (PCCA) was used to determine whether the effect of the contaminant PCA axis occurred separately to those of mud content or whether the effect of contaminants differed depending on how muddy a site was (as for example per Thrush et al (2008)).

3.2 Results

3.2.1 Development of the CAP Model

Analysis of the correlations between biological and sediment grain size data suggested that the mud fraction best explained species patterns (Table 1).

Table 1:

Relationship between the similarity matrix obtained using species data versus the Euclidean distance matrix obtained from the sediment grain size where 1= coarse sand, 2 = medium sand, 3 = fine sand and 4 = mud fraction.

Number of Variables	Correlation	Selection
1	0.39	4
2	0.39	1, 4
2	0.35	3, 4
3	0.35	1, 3, 4
4	0.33	All
3	0.33	2, 3, 4
2	0.33	2, 4
3	0.33	1, 2, 4
3	0.25	1, 2, 3
2	0.25	2, 3

3.2.2 Canonical Analysis of Principal Coordinates Results

CAP analysis based on sediment mud content resulted in a canonical correlation of 0.86 where m (the number of PCO axes used for the analysis) was equal to 14 and the proportion of the total variation in the dissimilarity matrix explained by the first PCO axes was 0.90 (Table 2). Table 2 also provides the results for the initial CAP model developed by Anderson et al., (2006) based on pollution gradients from PCA on log metal concentrations taken from the total sediment sample (PC1.500). Canonical correlations based on metal concentrations were slightly higher (0.89 vs 0.86 and the proportion of the total variation explained by the first PCO axes were also higher (0.92 vs 0.90.

Table 2:

Summary of CAP analyses to model sediment grain size using faunal data. Results from Anderson et al., (2006) for CAP analyses obtained from PCA on log metal concentrations on the basis of two different dissimilarity measures for the total sediment sample are also provided (PC1.500). m = the number of PCO axes used for the analysis, Prop.G = the proportion of the total variation in the dissimilarity matrix explained by the first m PCO axes, SSRES = the leave-one-out residual sum of squares, 1 is the squared canonical correlation for the canonical axis, Correl = the correlation between the canonical axis and the sediment or pollution gradient.

	Model	M	Prop.G	SSRES	1	Correl
Mud	Euclidean	14	0.90	0.385	0.74	0.86
PC1.500	Euclidean	16	0.92	0.342	0.78	0.89

3.2.3 Potential for interactions between the two separately derived CAP axes

The potential for interactions between the separately derived CAP axes (i.e., mud versus metal contaminants) was explored in three ways.

Firstly, a Pearson's correlation was run between the CAP scores along the Mud and the Metal axes. Correlation analysis of CAP scores for mud content versus CAP scores for metals indicated a moderate correlation (Pearson's R = 0.748). However while there is a moderate correlation, the relationship shows a high degree of variation (Figure 1). For example a mud CAP score of 0.5 could be associated with a metal CAP score ranging between -0.1 and 1.8. Conversely a metal CAP score of 0 could be associated with a mud CAP score ranging between -0.08 and 0.1. This indicates a considerable amount of variation explained by the varying scores along separate axes.

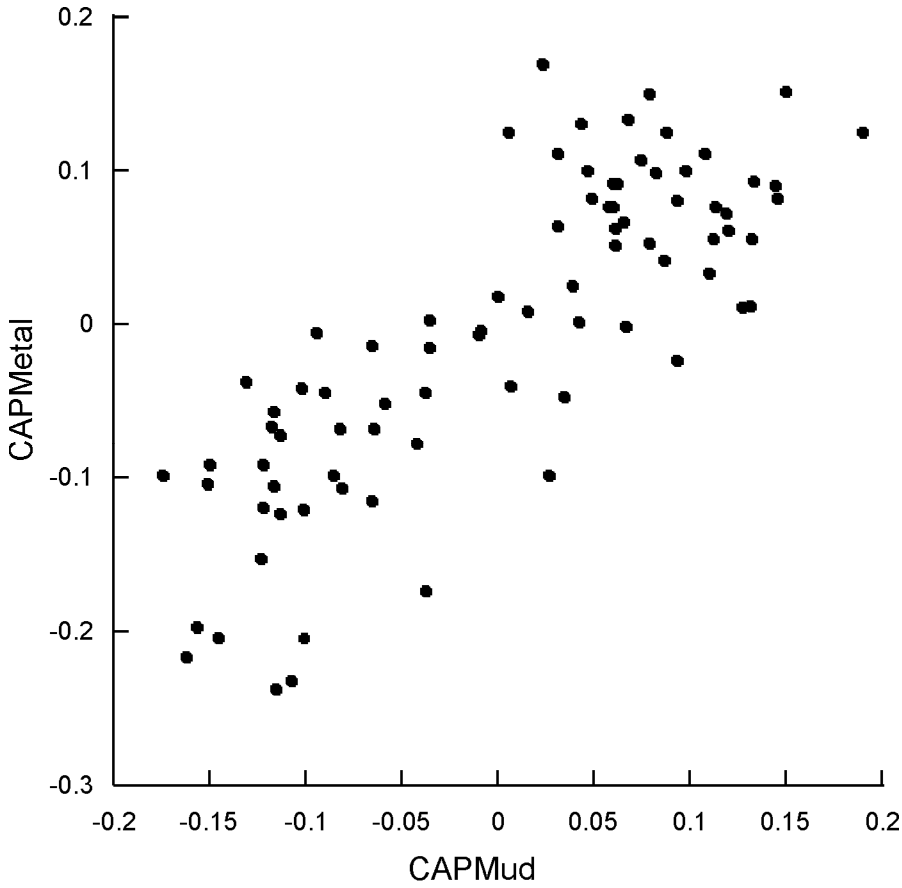


Figure 1:
Correlation between CAP axes related to mud content (CAPMud) and PC1.500 (CAPMetal).

Secondly, BIOENV was run to determine the best predictor variables based on both sediment grain size and PC1.500. Both mud and PC1.500 were selected as producing the best model (Spearmans ρ = 0.421.

Thirdly, there was only a 2% decrease in the amount of variability explained using mud content in a canonical correspondence analysis with the effect of PC1.500 partialled out, than there was prior to partialling out this effect.

3.3 Summary

A strong gradient of community change was observed in response to mud content of the sediment (correlation=0.86). This indicates that the Benthic Health Model can be used to determine potential effects of changes in sediment mud content. While there is a moderate correlation between site scores along the mud and metal axes, we feel that effects of sediment mud content can best be determined using a separate model for the following reasons.

- There was a high degree of variation in site scores along the CAP mud and CAP metal axes.

- Both Spearmans correlation and PCCA suggest that both mud content and metal contamination are important factors with neither being a replacement for the other.
- Anderson et al., (2006) investigated whether environmental characteristics including sediment grain size and wave exposure affected the relationship between the response of community composition to contamination and concluded (a) that the biotic assemblages had the strongest relationships with metal concentrations in the total sediment sample (PC1.500), rather than in the mud fraction (PC1.63), and (b) that the best model was one which included sites with both high and low mud content, rather than having separate models which discriminated effects based on mud content. This suggests that, despite muddier sites generally being more contaminated, there is still an effect of mud that is not related to contamination that we can determine using the models
- There is considerable information available on species sensitivities to increases in sediment mud content (Gibbs and Hewitt, 2004). Comparison between these sensitivities and the little information on sensitivities to heavy metal contamination (Hewitt et al., 2009, Thrush et al., 2008) suggest it should be possible to discriminate between the two stressors.

However, the correlation between the two sets of CAP scores does suggest that, when the health of a site is being assessed relative to changes in sediment characteristics and contaminant levels, a bivariate plot will be useful. Changes related to one axis would suggest a response to changes in that variable, while changes in both directions may be a result of responses to both factors and initiate a close inspection of which species are showing changes.

4 Testing the relationship between size class data and contaminant PCA

Recruitment events that occur prior to sampling may cause incorrect site assignments in the BHM. If post-settlement dispersal is causing problems with the model we would expect to see correlations between adults (and possibly large juveniles) and the contaminant PCA but not between smaller juveniles and the contaminant PCA. Therefore in this section we conducted correlations between various size classes of bivalves and contaminants. If there was a positive relationship with adults only and there was sufficient size-class data, we wished to build a new model to explore how sites change position with juveniles removed from the model.

4.1 Method

Size class information was available for the bivalves *Austrovenus stutchburyi and Macomona liliana* for 25 sites, a mix of model and validation sites (Table 3). The analysis is not affected by data from both model and validation sites because we are correlating abundances with the PC1.500 axis only. At each of these sites bivalves have been measured and categorized into four size classes. A Pearson's correlation was run for each size class/species with the contaminant PCA (PC1.500 Metals) to determine if they were correlated.

Table 3:

BHM sites for which size class information for bivalve species was available, including whether it was originally a site used to generate the model (M) or to validate it (V).

Site	Model / Validation
Benghazi	M
Bowden	M
Chelsea	M
Coxes	M
Hellyers Outer	M
Henderson Entrance (2002)	M
Henderson Entrance (2004)	M
Henderson Lower	M
Herald Island North	M
Hillsborough	M
Hobson Awatea	M
Hobson Victoria	M
Hobsonville	M
Main Outer	M
Meola Outer	M
Meola Reef (2002)	M
Meola Reef (2005)	V
Otahuhu	M
Panmure	M
Princes	M
Purewa	M
Shoal Hillcrest	M
Shoal Upper	M
Te Wharau	M
Whau Entrance	M

4.2 Results

Pearson's correlations indicate that there was a significant negative relationship between adult bivalves and the sediment metal contaminants (PC1.500) (Table 4). However this relationship was not present for juveniles of either species. For *Austrovenus stutchburyi*, significant negative correlations occurred for the 5-20mm size class. The larger size class of >20mm for *Austrovenus* had a reasonable p-value ($p = 0.11$), but this was not significant at the $\alpha = 0.05$ level probably due to the low numbers recorded in this size range. *Macomona liliana* showed significant negative relationships with metal contaminants for the size class range of >20mm. Moreover, the negative relationship with metal contaminants was stronger for *Macomona*.

Table 4:

Squared pearson correlations between differing size classes of the bivalve species Austrovenus stutchburyi and Macomona liliana and PC1.500.

Species	<1mm	<5mm	5-20mm	>20mm
Austrovenus	r^2 = -0.1962	r^2= -0.2386	r^2 = -0.4091	r^2 = -0.3197
	p = 0.3472	p = 0.2507	p = 0.0423	p = 0.1119
Macomona	r^2= 0.0835	r^2 = -0.0260	r^2 = -0.1668	r^2 = -0.5263
	p = 0.6912	p = 0.9018	p = 0.4255	p = 0.0069

4.3 Summary

The lack of any correlation between PC1.500 and abundances of juveniles of either *Macomona liliana* or *Austrovenus stutchburyi* suggests that the ability of juveniles to disperse in the water column may affect the Benthic Health Model's accuracy in discriminating between sites if recent dispersal has occurred. While data collection for the Benthic Health Model is limited to a time of year least likely for recruitment to occur, post-settlement juveniles can also disperse large distances throughout at least the first year of their life and *Austrovenus* often exhibits more than one recruitment period per year.

The problem generated by dispersal could be overcome by the development of a model in which juveniles and adults of these two species were treated as separate species rather than aggregated to a single count for each species. Unfortunately, there is currently insufficient data to produce such a model. We recommend that future monitoring routinely record size class information for bivalve species. This will slowly build up sufficient information to enable juveniles to be separated from adults for modeling purposes. In the interim it should be recorded whether individuals are very small and obviously newly settled or not and if potentially newly settled this site should be resampled.

5 Model sensitivity

While the Benthic Health Model was designed to assess health at new sites (those not in the original model), a major goal was to use the model to assess changes at previously assessed sites over time. To this end, a program of data collection was instituted by the ARC and data collection has been rotated around differing sites between 2006 to 2009. However, to date there have been no assessments of four points critical to the success of such monitoring.

- What degree of change can be detected by the model? Tricklebank et al (2010) note that changes in CAP scores of the model sites can occur when a new site is added in for assessment, raising the question of model sensitivity.
- What is the effect of adding differing numbers of sites into the model at once? This also allows us to assess how easy it would be to update or extend the model with new sites.
- What is the effect of adding new times into the model? This would allow us to determine whether changes over time must be assessed by adding a single site/time separately, or whether a single site with multiple times can be assessed simultaneously.
- What is the effect of new taxa being added to the model? Over time, even without the potential for invasive species to occur, new taxa are likely to be sampled due to the positive relationship between species accumulation and sampling effort (Arrhenius 1992, Hewitt et al., 1992, Gray 2002).

Since the model was completed and verified, sample collection at many of the validation sites has been ongoing. The effect of adding new sites, new times, new species and the sensitivity of the model to detect changes in community composition associated with increasing or worsening health are subsequently discussed below.

5.1 What degree of change can be detected by the model?

5.1.1 Methods

In order to assess the ability of the BHM to detect changes in species composition, we investigated change by simulating increasing or decreasing health of 10, 25 and 50% respectively. These levels were chosen as representing a range of percent change generally considered by ecologists. Ten percent is considered to be a weak change which may still have important consequences over the long-term, fifty % is considered to be a change likely to have dramatic ecological consequences.

Two model sites were selected from each of pollution rank indices (Table 5). Species were grouped into six categories based on their sensitivity or tolerance to increasing pollution (Appendix 1), or whether they were rare (as the number of rare taxa at a site has been observed to decrease with increasing contamination (Hewitt et al., 2009). For example species identified as sensitive to increasing pollution included *Anthopleura aureoradiata*, *Aonides trifida*, *Prionospio aucklandica*, *Colurostylis* sp., *Macroclymenella stewartensis*, *Macomona liliana*, *Paphies australis* and *Austrovenus stutchburyi* . Species that were tolerant to pollution or increased with contamination included the gastropod *Amphibola crenulata*, the spionid polychaete *Scolecolepides benhami*, crabs and Nereids. The species were determined using results from Anderson et al (2006), Thrush et al., (2003) & (2008); Ellis et al., (2006) and Hewitt et al., (2009).

Table 5:

Pollution index (where 1 = health and 5 = polluted) and site number used to test the BHM sensitivity to changes in composition.

Pollution Index	Site Number / Name
1	10/ Coxes 2004
1	35/ Mangemangeroa B
2	25/ Hobson - Tohunga
2	26/ Hobsonville 2002
3	51/ Motions East
3	79/ Turanga J
4	4/ Awatea 2004
4	53/ Ngataringa Bay
5	42/ Meola Inner 2005
5	92/ Whau Upper

The six categories of species (see Appendix 1) were then treated as follows:

1. Always decreased in abundance with increasing contamination- these species were multiplied by 0.9, 0.75 and 0.5 respectively for increasing contamination to create a change of approximately 10, 25 and 50% and similarly multiplied by 1.1, 1.25 and 1.5 for decreasing contamination.
2. Reached maximum abundance in Group 3. These were multiplied by 1.1, 1.25 and 1.5 for increasing contamination and by 0.9, 0.75 and 0.5 respectively for decreasing contamination for sites in Groups 1 and 2. Sites in groups 3 -5 were treated in the opposite way.
3. Reached maximum abundance in Group 4. These were multiplied by 1.1, 1.25 and 1.5 for increasing contamination and by 0.9, 0.75 and 0.5 respectively for decreasing contamination for sites in Groups 1 - 3. Sites in Groups 4 and 5 were treated in the opposite way.
4. Always increased in abundance with increasing contamination. These were multiplied by 1.1, 1.25 and 1.5 for increasing contamination and by 0.9, 0.75 and 0.5 respectively for decreasing contamination.

5. Were infrequently found. For decreasing contamination 10, 25 and 50% of these species were randomly allocated to a site and for increasing contamination 10, 25 and 50% of these species were randomly removed from a site.
6. Showed little effect. Abundances of these species were untouched.

The actual differences in community similarity between the initial and simulated communities were then calculated and represented averages of 5 (minimum to maximum of 3 – 5), 10 (7 – 11) and 17% (14 – 20). These actual changes in community achieved will be used in the text from this point on. As these changes were meant to reflect changes in contamination at the sites, the PC1.500 score of each simulated community was also altered (by either increasing or decreasing by 5, 10 and 17% to match the community data). The adjusted species abundance of each site was then added separately (so that only one site with one change was added to the model data for analysis) and a new CAP analysis was performed for each addition. The resultant CAP scores were then converted to run in an increasing direction with the PCA axis (if necessary), and to run over the same range as the CAP scores with only the model data in it (original scores range 0.403). The difference in the CAP score between the new positions of the model sites and the original scores were calculated as a % of the original CAP score range.

5.1.2 Results

The degree of change in CAP scores associated with the simulated increasing or decreasing health are given in Table 6. Firstly for the actual simulated increases, the average % change in CAP scores was -3.9, -7.3 and -12.6%. The maximum % change recorded ranged from -6.2 to -15.2 for a 17% increase in health. These compare to changes in CAP scores of the unmodified community composition associated with placing an extra site into the model of an average change of <1% and a maximum of 2.3%. The average change for 5, 10 and 17% decreases in health were 1.4, 4.0 and 8.1%. The maximum change for decreases in health of 17% ranged from 3.7 to 10.5%.

Table 6:

Maximum, minimum and average % change in CAP scores for sites with changed and unchanged community compositions, Sum>5 represents the number that exhibited changes of greater than 5%, maximum possible number is 8.

	Maximum	Minimum	Average	Sum > 5
Unchanged community	2.33	0.00	0.17	0
Increase 5%	6.20	2.96	3.88	1
Increase 10%	10.07	5.74	7.27	8
Increase 17%	15.22	10.55	12.61	8
Decrease 5%	3.66	0.51	1.43	0
Decrease 10%	6.28	1.70	3.99	3
Decrease 17%	10.45	5.60	8.14	8

These results suggest that the BHM is able to detect changes in health of communities, beyond the variability associated with the unchanged community.

Change in CAP scores associated with increases in health of only 5% can be observed (Figure 2) and are higher than variations in the scores of the unchanged communities (Table 6). Changes of 10 and 17% caused proportionally greater changes in the associated CAP scores. Small changes in decreased health were less easily detected. For the 5% decrease in health the minimum and average changes are less than the maximum variability observed for the unchanged communities. However, decreases of 10 and 17% were higher than the background variability.

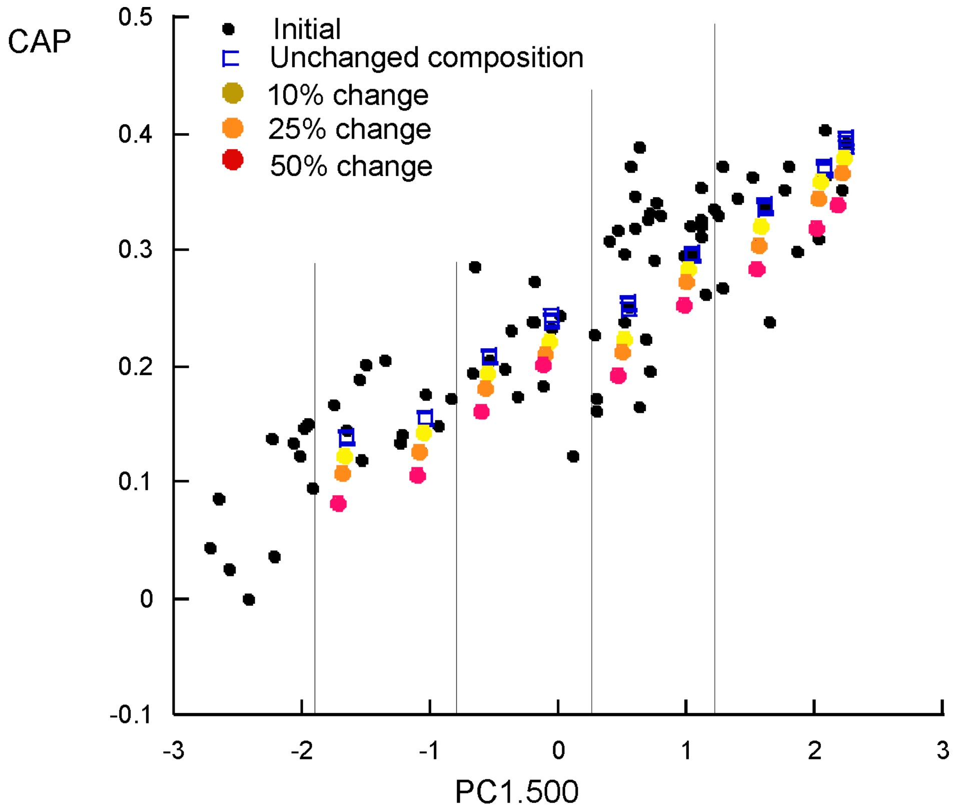


Figure 2:
The effect of increasing health by 10%, 25% and 50% on PCA and CAP scores for a range of sites in the BHM.

5.1.3 Summary

The BHM is capable of detecting changes in community composition as low as 5% for increasing health and 10% for decreasing health. These changes are detectable only by comparing CAP scores as such small changes in the PC1.500 are unlikely to result in sites crossing from one health category to another.

5.2 What is the effect of adding differing numbers of sites into the model?

5.2.1 Methods

The original models developed by Anderson et al., (2002) used 81 samples (M), with 14 samples being reserved to provide independent model validation (V). These validation sites can be used to test the effect of introducing "new sites" into the BHM (Table 7). Random selections of different numbers of extra sites were made from the validation sites in a stratified manner, to ensure that extra sites represented a gradient in effect. Five sets of 3, 6, 9, and 12 sites were selected, and a CAP was performed for each of these random selections. The resultant CAP scores were then converted to run in an increasing direction with the PCA axis, and to run over the same range as the CAP scores with only the model data in it (original scores 403). Finally the difference in the CAP score between the new positions of the model sites and the original scores were calculated as a % of the original CAP score range. The maximum, minimum and average percent change of the new positions of the model sites relative to the original scores were calculated. The number of sites that exhibited changes of greater than 5% were also calculated. The correlation with the PCA axis was also recalculated each time a new analysis was run, but the results indicated little change in the correlations. Finally, all 14 validation sites were added into the model and the same analyses and statistics were performed.

Table 7:

Validation sites used in the analysis.

Auckland Airport
Coxs
Henderson Upper 2005
Herald Island
Hobson - Purewa Bridge
Kendalls
Little Shoal Bay
Mangemangeroa E
Mangere Inlet: Tararata Creek
Newmarket
Okura J
Pakuranga Mid
Whau Entrance, WHO A
Whau Lower

5.2.2 Results

Results of the average % change in the CAP score between the new positions of the model sites and the original scores ranged from 0.6% to 1.9% (Table 8). The maximum change was 12.0% when 12 sites were added into the model run. Finally the number of sites that exhibited a change of greater than 5% ranged from 11 to 17 sites (Table 8).

Table 8:

Results of the maximum, minimum and average % change between the new positions of the model sites and the original PCA scores as additional sites are added into the BHM. Sum>5 represents the number of sites that exhibited changes of greater than 5%.

Number of sites added	Maximum	Minimum	Average	Sum > 5
3	8.47	0.00	0.60	11.0
6	9.67	0.00	0.70	13.3
9	10.26	0.00	0.79	16.5
12	12.00	0.00	1.90	16.3
14	9.00	0.00	0.90	17.0

5.2.3 Summary

While the average % changes that occurred with the addition of new sites were small, the number of sites that exhibited changes of greater than 5% was quite high. These changes in position suggest that assessment of sites should be done on a single site at a time as recommended by Anderson et al., (2006). However, the lack of change in correlation between the CAP and PC1.500 axis suggest that new sites can be used to update the model without weakening the model. This should only be done after due consideration and recalculation of previous results.

5.3 What is the effect of adding new times into the model?

5.3.1 Methods

From the new data collected between 2006-2009 there were three new sites (Kendalls, Newmarket and Pakuranga) that were sampled in multiple years. By using these three sites we can test the effect of adding new times versus new sites into the model. All of these sites are Validation sites that were collected when the original Model was developed (Time 1). Further all sites were sampled an additional 2 times when the new data were collected between 2006-2009 (Time 2 and Time 3).

New PCA values were created by using the conversion factors given in Anderson et al., (2006). Specifically the following equation can be used to place a new site on the PC axis based on its metal concentrations;

PC1.500 = $0.615\times(X_{Cu}^{(500)})+0.528\times(X_{Zn}^{(500)})+0.586\times(X_{Pb}^{(500)})$

where, for example, X_{Cu}= the log concentration of copper in the total sample (<500 m) minus the mean log concentration of copper (<500 m) across the full set of 81 samples, and so on for the other variables. The mean log concentrations for copper, zinc and lead are provided in Anderson et al., (2006).

Three separate CAP analyses were then run. The analysis was performed in a step-wise fashion in order to determine the effect of adding new times and new sites into the model. Specifically the 3 validation sites collected at Time 1 were added into the model and the degree of change of all model sites calculated as per section 5.2.1. Then Time 2 was added into the model and again the degree of change of all sites/times sites calculated. Finally Time 3 was added into the model and the degrees of change calculated.

5.3.2 Results

Results indicate the average % change on model sites was slightly less when new times are added into the BHM (Table 9). An average change of 0.3 occurs when the 3 new sites are added into the model, a similar amount occurs when the extra times at the 3 sites is added (0.6 – 0.3) and finally an average (0.1) change occurs when the next time is added (0.7 – 0.6). When comparing the % change at the new sites alone, the average % change increased from 3.9% with the addition of 1 new time to 4.4% with the addition of 2 new times (Table 9).

Table 9:

Results of the maximum, minimum and average % change between the new positions of the model sites and the original PCA scores as additional sites and times are added into the BHM.

	Maximum	Minimum	Average
Change at model sites			
Add 3 new sites	7.83	0	0.28
Add 3 new sites at 1 time	5.15	0	0.65
Add 3 new sites at 2 times	5.67	0	0.66
Change at 3 new sites only			
With 1 new time	5.43	0	3.90
With 2 new times	9.65	0	4.41

5.3.3 Summary

Overall adding new times had a lesser effect on the model than did the addition of new sites into the model. Based on the results above we recommend that for existing model sites, new times can be assessed simultaneously, but when new sites are added changes should be assessed by adding a single new site/time separately.

5.4 What is the effect of adding new species into the model?

5.4.1 Methods

Analysis of the species identified in the new datasets collected from 2006 to 2009 identified a total of eight new species. These species included *Pomatoceros* sp., *Carazziella phillipensis*, *Potamopurgus* sp., *Cantharidus purpureus*, *Cominella lineolata*, *Lasaea parengarenga*, *Neoguraleus interruptus* and *Parawaldeckia* sp. These species were sampled at Meola Inner, Shoal Bay Lower, Whakataka, Newmarket, Chelsea, Coxes and Panmure sites and are used to investigate the effect of adding new species into the BHM.

CAP analyses was run without the new species and then with the new species and degree of change calculated as per section 5.2.1.

5.4.2 Results

The maximum % change was 0.2% with the addition of the new species. There were no changes >5% and the average % change was approximately 0%. These results indicate little effect of the addition of new species on the BHM.

5.4.3 Summary

At present, only taxa present in the initial 81 sites are allowed to influence community composition with new species being omitted from the model. However, over time new taxa are likely to be sampled as a result of the positive relationship between species accumulation and sampling effort (Thrush et al., 1988; Hewitt et al., 1992). The results indicate that there was relatively little effect on the BHM of the addition of new species into the model.

6 Summary and Recommendations

6.1 Development of the BHM for State of the Environment monitoring

Extending the present BHM from its focus on storm-water pollution to a State of the Environment tool requires the determination of other anthropogenic activities that are likely to affect benthic communities. A meeting with the Auckland Regional Council was held in January 2010 to identify key potential anthropogenic disturbances for possible inclusion in the model. The discussion focused on three sources of anthropogenic stress including; 1) disturbances associated with changes in sediment grain size, 2) changes in other chemical contaminants not currently modeled by the BHM, and 3) confounding variables.

This research using available sediment grain size data was able to clearly demonstrate strong gradients of community change in response to mud content of the sediment using the BHM. While moderate correlations between site scores along the mud and metal axes were found, the use of bivariate plots of CAPMud and CAPMetal axes could be used to assess site changes and hence species responses to the two varying anthropogenic stressors (Figure 3). For example, movement of a site along one axis, say the CAPMetal axis, should be investigated for species changes in response to metal contaminants. Movement along the CAPMud axis would indicate species responses to elevated mud content. Site changes in both directions may be a result of responses to both factors and would initiate a close inspection of which species are responding to the changes.

These determinations could also be supplemented by literature that currently provides species sensitivity information to changes in sediment mud content. Considerable information is currently available including a review of the primary reports and scientific papers by Gibbs and Hewitt (2004) that provides a list of species responses to changes in sedimentation rates. Specifically the sediment preferences, distribution, and optimal ranges of common macrofaunal species are documented. Species-specific models that predict probability of occurrence have also been developed (Gibbs and Hewitt, 2004; Thrush et al., 2003; Ellis et al., 2006). Therefore if a site is found to be moving along the mud axis these models can directly be used to determine likely species responses to changes in the sediment regime that can then be assessed with the site specific data. If a site moves along the PC1.500 axis species are likely responding to metal contaminants and this information over time can be used to increase our understanding of toxicity effects of benthic macrofauna which is currently limited.

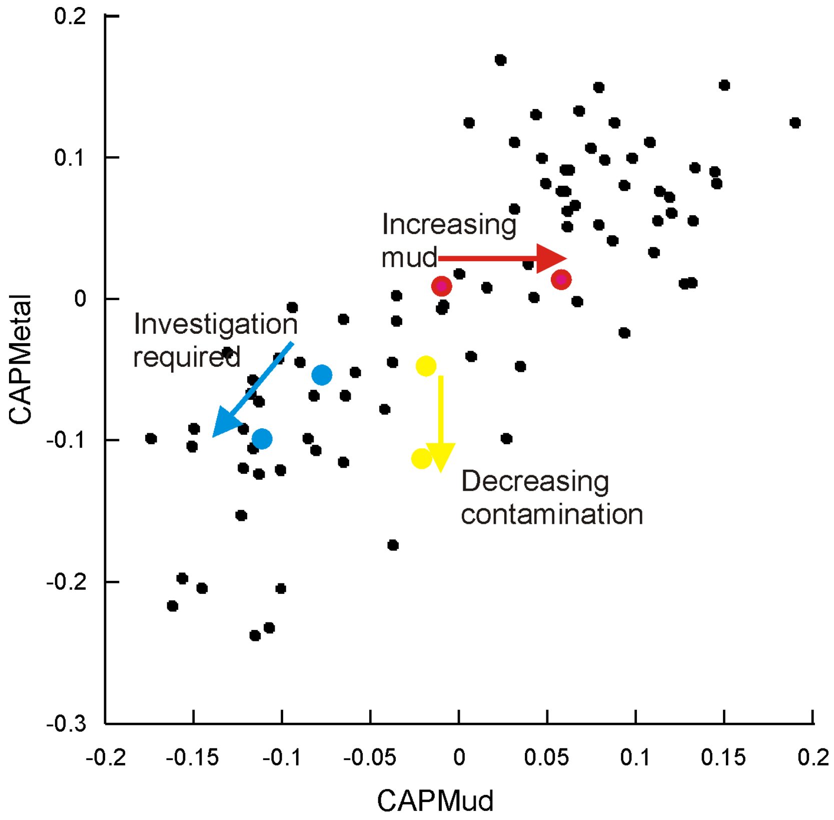


Figure 3:
Use of bivariate plots to in SOE monitoring.

Chemical contaminants that are not currently modeled by the BHM identified by the ARC for consideration included TPHs, PAHs, organochlorines, ammonia and dissolved chemicals present in the water. The ARC determined that there is currently inadequate data on TPHs and limited data for PAHs presently for modeling purposes. Again there is presently no information on dissolved levels of copper, zinc and lead in the water on which to assess their effects on benthic communities. TPHs are likely to be included in the new ANZAC guidelines and while there is currently inadequate information, could be considered for future monitoring. As adequate spatial data is obtained these parameters could then be tested for their correlation with the stormwater PC1.500 axis. If they are not highly correlated they could then be modeled as an extra contamination gradient into the BHM. Currently there is sufficient spatial data for some contaminants (e.g., ammonium) to assess whether other contaminants are causing obvious outliers from the benthic health mud content and stormwater contaminant models.

Finally, confounding variables were identified that included the geology of an area, the bioavailability of a given metal, and the organic carbon and nitrogen in the sediment. These factors could be incorporated into future models as covariables where adequate data is available and again are recommended for consideration in future monitoring programs.

Differences in response time of different life stages of bivalves driven by dispersal were also identified as a potential confounding influence on the BHM. At present, dealing with the problem generated by dispersal can be dealt with by investigating obvious outliers in response. At present there is insufficient data to allow development of a model in which juveniles and adults of these two species were treated as separate species rather than aggregated to a single count for each species. If future monitoring routinely records size class information for bivalve species a new BHM could be developed in the future that was less susceptible to dispersal of juveniles.

6.2 Sensitivity of the Benthic Health Model

The BHM was able to detect most of the simulated changes. Model increases or decreases in health showed corresponding changes between the new community composition and the original scores related to the level of simulated change. The model demonstrated a slight increased ability to detect movement towards health (5% change in community composition detected) rather than that of worsening health (only changes > 10% able to be detected). No movement between rankings was observed, with the model providing a more sensitive assessment of change if actual movement along the CAP axis is calculated. This is because whether a site changes rank is highly dependent on how close it is to a boundary between rankings with the PC1.500 rankings covering approximately a 20% change in values.

Adding new sites into the model had little effect, though the number of sites that exhibited changes of greater than 5% was quite high. Assessment of new sites should be done by adding new sites to the model one at a time. However, the lack of change in correlation between the CAP and PC1.500 axis suggest that new sites can be used to update the model without weakening the model. However, for existing model sites, new times can be assessed simultaneously.

There was relatively little effect on the BHM of the addition of new species into the model. Part of this is potentially due to the fact that the BHM has broad taxonomic groupings such as Goniadidae, Nereididae, Oligochaetes, Other Amphipods, Crabs, and so forth. These general groupings mean that species that may not have been identified in the Model development but which fall under a general taxonomic grouping will be treated in the same categories when new times are added. It should be noted that the new species identified were not highly numerically abundant. We recommend that in the case of new species that are highly abundant or invasive the effect on the BHM would need to be reassessed.

7 Acknowledgements

Assistance for this report has been received from Dr Marti Anderson (Massey University), Dr Chris Hickey (NIWA), Mr Andrew Swales (NIWA), Dr Geoff Mills (Diffuse Sources Ltd) and Drs Megan Stewart and Jarrod Walker (ARC).

8 References

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

Anderson, M.J.; Hewitt, J.E.; Ford, R.B.; Thrush, S.F. (2006). Regional models of benthic ecosystem health: predicting pollution gradients from biological data. Prepared by Auckland UniServices Ltd for Auckland Regional Council. *Auckland Regional Council Technical Publication 317.*

Auckland Regional Council (2002). Environmental targets for the urban coastal marine area. *TP169*. 26 p.

Arrhenius, O. (1921). Species and area. *Journal of Ecology 9*: 95–99.

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

Gibbs, M.; Hewitt, J. (2004). Effects of sedimentation on macrofaunal communities. A synthesis of research studies for ARC. *Auckland Regional Council Technical Publication 264.* 54 p.

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

Hewitt, J.E.; McBride, G.B.; Pridmore, R.D.; Thrush, S.F. (1992). Patchy distributions: optimizing sample size. *Environmental Monitoring and Assessment 27*: 95–105.

Hewitt, J.E.; Anderson, M.J.; Thrush, S.F. (2005). Assessing and monitoring ecological community health in marine systems. *Ecological Applications 15(3)*: 942–953.

Hewitt, J.E.; Anderson, M.J.; Hickey, C.W.; Kelly, S.; Thrush, S.F. (2009). Enhancing the ecological significance of sediment contamination guidelines through integration with community analysis. *Environmental Science and Technology, 43*: 2118–2123.

Thrush, S. F.; Pridmore, R.D.; Hewitt, J.E.; Roper, D.S. (1988). Design of an ecological monitoring programme for the Manukau Harbour. Unpublished report for the Auckland Regional Water Board - Auckland Regional Authority. *Water Quality Centre Consultancy Report 7099*.

Thrush, S.F.; Hewitt, J.E.; Norkko, A.; Nicholls, P.E.; Funnell, G.A.; Ellis, J.I. (2003). Habitat change in estuaries: predicting broad-scale responses of intertidal macrofauna. *Marine Ecology Progress Series, 263*: 101-112.

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.

9 Appendix 1

Taxa sensitivities to stormwater contamination. Freq occur = frequency of occurrence in model data.

	Taxa	Freq occur
Group 1	Always decrease in abundance with increasing contamination	
	Anthopleura aureoradiata	25
	Aonides trifida	25
	Prionospiio aucklandica	64
	Austrovenus stutchburyi	62
	Barnacles	20
	Colurostylis spp.	38
	Exogoninae	32
	Macomona liliana	48
	Macroclymenella stewartensis	22
	Orbinids	38
	Paphies australis	19
	Scolelepis spp.	9
	Waitangi brevirostris	9
Group 2	Reach maximum abundance in Group 3	
	Capitella + Oligochaetes	53
	Cominella glandiformis	33
	Euchone sp.	12
	Glycera spp.	45
	Goniadidae	18
	Notoacmea spp.	27
	Nucula hartvigiana	61
	Tanaidacea	10
	Zediloma subrostrata	14
Group 3	Reach maximum abundance in Group 4	
	Aricidea sp.	47
	Cossura consimilis	43
	Heteromastus filiformis	75
	Nemertean	74
	Polydorid complex	77
	Theora lubrica	31
	Zeacumantus lutulentis	14

	Taxa	Freq occur
Group 4	Always increase in abundance with increasing contamination	
	Amphibola crenata	29
	Arthritica bifurcata	68
	Helice, Hemigrapsus, Macropthalmus	75
	Nereidae	76
	Paracalliope spp.	50
	Paraonid	14
	Phoxocephalids	75
	Scolecolepides benhami	51
Group 5	Rare species	
	Amalda spp.	1
	Anthuridae spp.	1
	Asychis amphiglypta	2
	Bulla quoyi	2
	Carditidae	1
	Cirolana sp.	2
	Diastylopsis sp. (Cumacea)	1
	Felaniella zelandica	2
	Harmothoe sp.	1
	Lumbrinereidae	1
	Maldanidae	0
	Mantis shrimp	1
	Minuspio	1
	Nebalace	3
	Opistobranch (Philine type)	1
	Owenia fusiformis	2
	Palaemon affinis	0
	Phyllodocid spp.	2
	Platyhelminth	2
	Pontophilus australis	2
	Scintillona zelandica	0
	Solemya parkinson	2
	Spionid	2
	Sipunculid	1
	Tellina edgari	1
	Travisa olens	0
	Trochotodota dendyi	1
	Venericardiae	2

	Taxa	Freq occur
	Xymene sp.	3
	Zegaluri tenius	3
Group 6	Others	
	Aglaophamus macroura	5
	Alpheus	20
	Amphipod	40
	Armandia maculata	10
	Bivalve unid	7
	Cirratulid	15
	Chiton	12
	Cominella adspersa	6
	Corophidae	50
	Crassostrea gigas	2
	Cyclaspis thomsoni	4
	Disconatus accolus	4
	Edwardsia	3
	Exosphaeroma spp.	19
	Gastropod unknown	4
	Halicarcinus spp.	35
	Haminoea zelandiae	7
	Hessionid	3
	Hiatula	9
	Isopod other	13
	Lepidonotinae	11
	Mactra ovata	32
	Magelona ident	12
	Micrelenchus sp.	5
	Musculista senhousia	4
	Mysidacea	20
	Notomastus	1
	Ophiuroid	4
	Paralepidonotus ampulliferus	7
	Pectinaria australis	19
	Phoronid	4
	Pinnotheres	12
	Polynoid	1
	Sabellidae	7
	Syllinae	4
	Turbonilla sp.	9