Auckland
Regional Council
TE RAUHITANGA TAIAO

Regional Models of Benthic Ecosystem Health

Predicting Pollution Gradients from Biological Data

November 2006 TP317

Auckland Regional Council
Technical Publication No. 317, 2006
ISSN 1175-205X
ISBN -13 : 978-1-877416-50-7
ISBN -10 : 1877416-50-9

Regional models of benthic ecosystem health: predicting pollution gradients from biological data

M. J. Anderson[1],
J. E. Hewitt[2],
R. B. Ford[1] and
S. F. Thrush[2]

Prepared for: **Auckland Regional Council, August 2006**

Prepared by: Auckland UniServices Limited
A wholly owned company of
The University of Auckland

[1]Department of Statistics & Leigh Marine Laboratory
University of Auckland, Private Bag 92019
Auckland, New Zealand
Tel: 64-9-373-7599 x85052
Fax: 64-9-373-7000
Email:mja@stat.auckland.ac.nz

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

1 TABLE OF CONTENTS

2 EXECUTIVE SUMMARY

The purpose of this work was to develop new regional models of benthic ecosystem health for sheltered intertidal soft-sediment habitats on the basis of new and existing biological, chemical and physical data. More particularly, we wished to obtain a model whereby biological data from a new or monitored site could be used to classify that site in terms of its relative health. Data assembled from sites across the Auckland Region included mean abundances of 102 taxa from 84 sites, some of which were sampled in multiple years (from 2002-2005), yielding 95 samples. Models were developed using 81 samples, with 14 samples being reserved to provide independent model validation. Physical data included grain size fractions and measures of furthest and closest wind exposure. Chemical data consisted of measures of concentrations (mg/kg) of copper, lead and zinc from the total sediment sample (< 500 μm) and also from weak acid extraction of the mud fraction (< 63 μm). The latter is generally considered a measure of bioavailable metals.

Metal concentrations showed very high correlations with one another, so a single measure of the degree of pollution along a gradient across all samples was obtained using principal components analysis (PCA). This was done separately for the total sediment measures (PC1.500) and for the mud fraction measures (PC1.63), explaining 94% and 95% of the variation in metal concentrations, respectively. If metal concentrations are sampled at a site, the degree of pollution can be determined directly by calculating the position of that new site, given these values, along each of these gradients. Clusters of 5 groups were identified along each gradient (in rank order from 1 = healthy to 5 = polluted). Groups 4 and 5 along the gradients corresponded well with existing "amber" and "red" sediment quality guidelines of the Environmental Response Criteria ("ERC") (ARC 2004). However, the PC axes developed here give greater resolution and discrimination among healthier sites (groups 1-3).

Ecological assemblages generally reflected pollution gradients very well, all along their range. The present study identified clear methods for modeling the pollution gradient axes using ecological data. Canonical analysis of principal coordinates (CAP) was used to develop models using biotic dissimilarities among sites to predict their relative position along each of the pollution gradients (PC axes). The best models of benthic ecosystem health were those which obtained high canonical correlations with the pollution gradient(s) and which had a low level of error when new sites were tested (validation).

The best overall ecological models were obtained using all sites together, regardless of their physical characteristics. 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), indicating that the biota do respond to all metals present in the sediments. Some of the models which used only subsets of taxa (a biologically derived "sensitivity" subset of 22 variables and a statistically derived "BVSTEP" subset of 16 variables) performed virtually as well as the models which used all 102 taxa. Although

we do not feel that models using subsets can replace those which use all taxa, they may be used with fairly high confidence if the data for all taxa, for some reason, is lacking.

The physical characteristics at each site (grain size fractions and exposure indices) were used to identify two physical groups of sites: those having coarser sediments and greater exposure (group C) and those having finer sediments and lesser exposure (group F). These two physical groupings correspond roughly to the Outer Zone and Settling Zone, respectively. Although no advantage was obtained by relating sites in group F alone to total metal concentrations (PC1.500), an excellent model was obtained by considering sites in group C alone and relating these to metal concentrations in the mud fraction (PC1.63). This supports previous studies suggesting that heavy metals are potentially more bioavailable in the mud fraction in Outer Zones.

We recommend that the models we have developed here be used for monitoring and management purposes, as follows:

- First, using all taxa, the position of a new site (or a monitored site) may be obtained along PC1.500 (and therefore into a group of relative pollution from 1-5) on the basis of the biotic dissimilarities between it and each of the existing sites, which we will call "the first classification".
- Second, if the site has relatively coarse sediments and greater relative exposure (i.e. if it occurs in group C based on its physical characteristics, or is in the Outer Zone), then its biotic dissimilarity with all other group C sites will yield its position along PC1.63 (and therefore, once again, into a group of relative pollution from 1-5), which we will call "the second classification".
- For sites in group C, the more cautious (i.e. the higher value) of the first and second classifications can be used as the assessment of the benthic health of the site. For sites in group F, the first classification can be used.
- If metal concentration data from the site is available, then validation of the positioning of the site on each of PC1.63 and PC1.500 achieved by using biotic data can be obtained. The ERC criteria can also be examined and considered when metal data are available.

A multivariate computer software package, PRIMER v6, with the add-on PERMANOVA+ (Anderson and Gorley – to be released in 2007), will be provided so that ARC managers can implement this strategy of modeling directly.

These models rely on the high degree of correlation in the levels of the three metal concentrations co-occurring in sediments across the region. It is anticipated that, even though zinc and lead, in particular, are currently highly correlated ($r = 0.95$), these associations may change with contaminant control measures. For example, lead levels are expected to decrease across the region, due to the elimination of leaded petrol. Zinc levels may decrease over time with controls on metal cladding and roofing. Thus, we further recommend that, from time to time, the degree of correlation among the three

metal variables (copper, lead and zinc) be checked. If these correlations begin to decrease, then a re-development of the overall model which separates and distinguishes the three individual metals may be appropriate.

3 BACKGROUND and PURPOSE

Auckland's Regional Discharge Project (RDP) is concerned with the management and treatment of stormwater and its effects on the biological quality of estuaries. Previous study and work to date has resulted in the development of quantitative models of ecosystem health (Anderson et al. 2002, ARC 2002, Hewitt et al. 2005). The purpose of the models 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 previous models of "health" relied, however, 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 other existing knowledge by ARC managers.

It was proposed that the benthic health model be refined and developed. More particularly, it was determined that an appropriate model should 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. In addition, extra sites were nominated as "validation sites" in order to explicitly and independently test the utility of the model(s) for active environmental management.

The purpose of the present work is:

1. to develop new regional models of benthic ecosystem health on the basis of new and existing biological, chemical and physical data;
2. to obtain independent validation of the models using data from additional validation sites and
3. to provide a recommendation demonstrating how benthic health can be assessed by ARC managers using new models.

4 SAMPLING METHODS

The selection of sites, the methods of sampling and the processing of samples were all consistent with the methodologies provided in ARC TP 168: "Blueprint for monitoring urban receiving environments" (ARC 2002).

4.1 Selection of sites

Eighty-four sheltered intertidal soft-sediment sites were selected for sampling within the Auckland region which together fulfilled the following four criteria:

- They covered the range of levels of pollutants present in these habitats;
- They covered the range of grain sizes present in these habitats;
- They were located on homogeneous unvegetated intertidal flats away from obvious discharge pipes, mangrove pneumatophores or low-tide channel banks; and
- They covered the probable geographic range of future large-scale urban development.

Data were drawn from several existing monitoring programmes (i.e., Kingett Mitchell Limited 2002, Funnell et al. 2003, Hewitt et al. 2004, Reed and Webster 2004, Williamson and Kelly 2003, Ford and Anderson 2005, Hewitt et al. 2006) and were also supplemented with targeted sampling of additional sites to complement existing information. Sites extended from Puhoi estuary in the north, to Clarkes Beach in the south (Fig. 1) and ranged in size from 100m × 50m to 100m × 100m. Each site was given a number for identification in the present study and all sites were sampled in October or November of 2002, 2004 or 2005. This time of year was chosen in order to minimise the potential effects of seasonal recruitment events on measures of biotic assemblages. At a few of the sites, samples of sediment texture were collected in May 2006, and these were matched with biological samples from those same sites which had been obtained in late 2005 (see Appendix 1 for details). This was considered reasonable, as long-term monitoring data from many of these sites indicate a relatively high degree of stability in recorded ambient sediment texture for purposes of broad-scale description (e.g., Ford and Anderson 2005).

The resulting set of data included information from some sites for more than one year of sampling. For these sites, samples were retained and included in analyses only if they had complete and independent information for all sets of variables of interest (i.e., biology, sediment texture and chemistry). This resulted in a total of 95 sampling units (each uniquely identifiable as a particular site in a particular year) available for analysis in the present study and a complete list of all of these is given in Appendix 1.

The process of modeling required two sets of samples: a modeling set and a validation set. We wished to use the maximum available information for the modeling set, yet still

to have some independent information from new sites in the region for validation of the models developed. The ARC (through its representative Dr Shane Kelly) were asked to provide a list of sites to be used for validation purposes. These were chosen according to the following criteria:

- They were spatially interspersed with modeling sites;
- They spanned the region of interest; and
- They did not contain the most extreme values of any of the chemical or physical measures, so that they would not fall outside the bounds of what had been developed using modeling sites

There were 14 validation sites selected as above, so 81 of the 95 samples remained for model development. These are indicated in Appendix 1 and Figure 1.

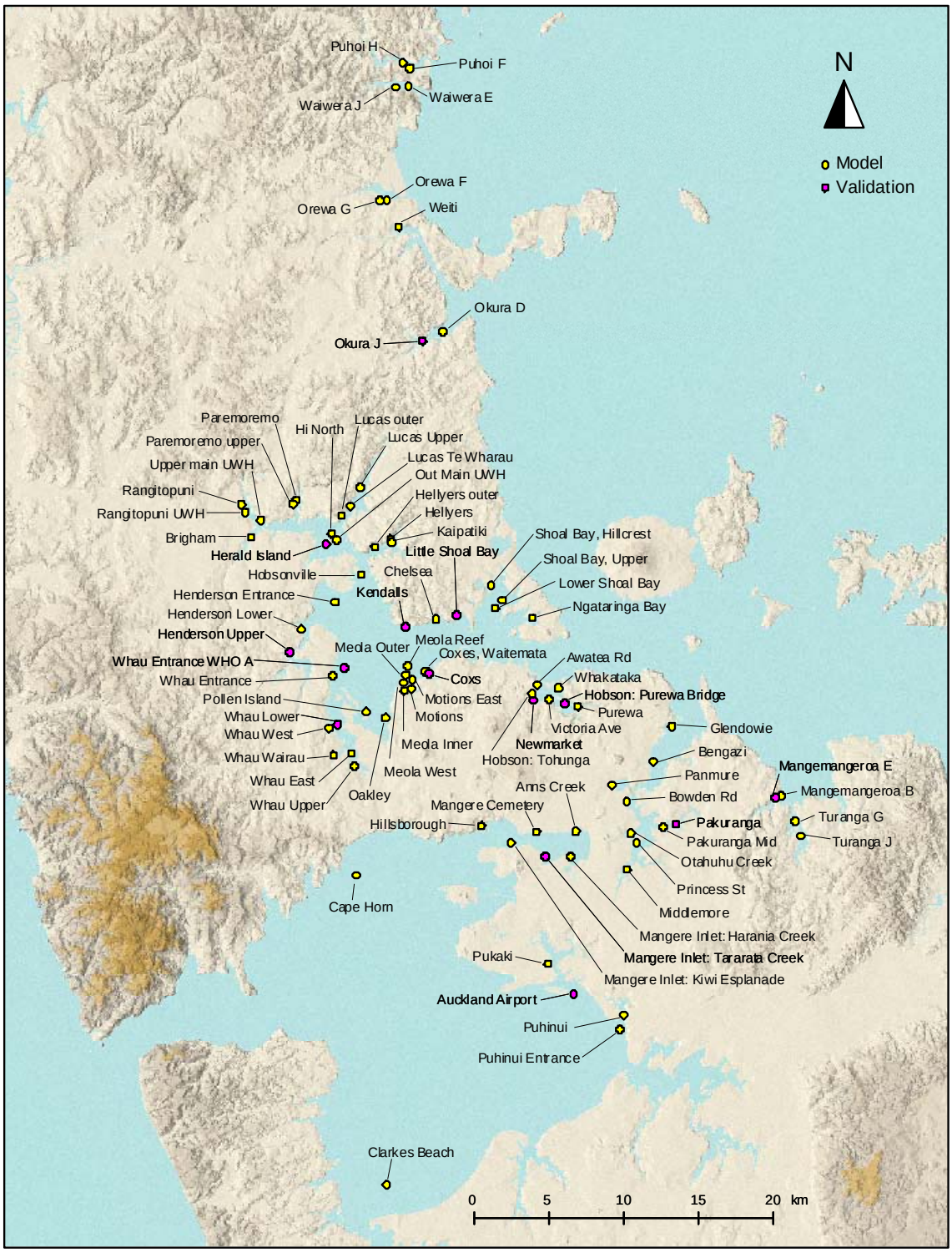


Fig. 1. Map of the Auckland region showing the 95 sites included in this investigation: 81 sites for developing models (in yellow) and 14 sites for model validation (in pink).

4.2 Sampling of the biota

At a given site and time, biological samples were taken using a circular corer measuring 130 mm in diameter by 150 mm deep. Depending on the particular source for the data, the number of cores per site varied between $n = 6$ and $n = 12$, with most sites having sample sizes of $n = 10$. Cores were taken from random positions within each site. Each core was sieved using 0.5 mm mesh and the material retained on the sieve was preserved in 70% isopropyl alcohol with 0.01% rose bengal. All organisms were identified in the laboratory to the lowest practical taxonomic resolution. Some discrepancies in the level of taxonomic resolution used in different studies meant that the lumping of certain groups of organisms was necessary in order to proceed with a single combined dataset. Although such lumping is regrettable, the resulting full biotic dataset nevertheless did retain a total of 102 separate unique taxa for analysis.

For all analyses relating the biota to environmental and chemical data, a single vector of abundance values across all taxa was required for each site and year of sampling. A list of the taxa and their frequencies of occurrence (out of 95 sampling units) is provided in Appendix 2. We used the vector of averages for each variable (i.e., the centroid) as an appropriate summary measure of the community structure at each site/time. Although the sample sizes differed (from $n = 6$ to 12, with most sites having $n = 10$), the degree of bias due to lack of detection of species for smaller sample sizes was considered negligible and there was no detectable bias in total average abundances for different sample sizes (see details in Appendix 3).

4.3 Measures of sediment texture and exposure

Samples of sediment texture were obtained using a core measuring 20 mm in diameter by 20 mm deep adjacent to each biological sample. The cores obtained from each site were combined prior to processing. Sediments were first digested in 6-10% hydrogen peroxide until bubbling ceased, then were momentarily boiled prior to wet-sieving. Sediments were wet-sieved through four sieves (2mm, 500µm, 250 µm and 63 µm) prior to decanting and drying of each fraction at 60°C. Five variables were used to describe sediment texture, consisting of the percentage composition by weight occurring in each of five classes: gravel (>2mm), coarse sand (2mm – 500 µm), medium sand (500µm - 250 µm), fine sand (250 µm – 63 µm) and silt and clay (<63 µm). Gravel and coarse sand were combined, resulting in four variables for subsequent analyses.

In addition to the sediment texture, the degree of exposure of individual sites was also measured using two variables. Distances to the nearest and furthest land point were measured for each of 8 compass sectors (i.e, 0° ± 22.5°, 45° ± 22.5°, and so on) for each site. Data from the Musik Point wind station from 2002 - 2005 were compiled into a frequency/strength histogram for each of the compass sectors. These distances were multiplied by the frequency/strength values for each compass sector and then summed. The values obtained for each of these two variables were scaled from 0-100

by taking them as a percentage of the largest value obtained across all sites. The two resulting variables are referred to hereafter as furthest wind exposure (FWE) and closest wind exposure (CWE). In the case of FWE, an effectively infinite distance for one of the fetch values resulted in an extremely large measure at one of the sites (site 31, Mangemangeroa B), which appeared as an outlier (with an FWE value of 100) compared to all of the other values in the dataset for that variable (which were all less than 19). This outlier was therefore given an arbitrary value of 20, retaining its ranking as the largest in the dataset, but without unduly affecting the nature of the distribution of values for that variable.

4.4 Sampling of heavy metals

At a given site and time, sediment samples were collected to a depth of 2 cm using a scoop made from a square unused polyethylene bottle for analysis of the heavy metals copper (Cu), lead (Pb) and zinc (Zn). In some cases, a sample was taken alongside each biological core and these were then combined to obtain a single composite representative sample. In most cases, however, metals were sampled separately in a 20 m × 50 m area adjacent to biologically sampled sites. There were three replicates per site, and each replicate was made up of 10 sub-samples taken every 2 m along the length of the site and assigned in a sequential manner to the three replicates (e.g., ARC 2004). Thus, each composited replicate contained sub-samples that covered the entire spatial area of the site. Metal concentrations for each replicate were assessed using two techniques: (i) weak acid digestion in 2 M HCl of the <63 μm (mud) fraction at room temperature and (ii) strong acid digestion of the freeze dried < 500 μm (total) sediment fraction in aqua regia (HCl/HNO_3) at 100-110°C. (See p. 20 of the Blueprint document (ARC 2004) for more details.) Mean values of the three replicates obtained for each site were then used in subsequent analyses.

The mud fraction (< 63 μm) is thought to be the most ecologically relevant component of sediments in terms of contaminants. The metal concentrations obtained from weak acid digestion of this fraction should broadly correlate with bioavailable amounts of contaminants. More particularly, this measure is intended to reflect the metals that would be available were fine sediments to be ingested and digested in a weak acid environment inside organisms. However, while the majority of particles ingested by macrofauna are < 63 μm, larger sized particles (up to 500 μm) may be ingested. Strong acid digestion of the whole sample will effectively extract all of the metals that are present and bound to the surfaces of sediment particles. This therefore provides a measure of the total amount of metal in those sediments, without presuming the nature or extent of the bioavailability of those metals. Either or both of these sets of measures could prove useful for assessing ecosystem health in these habitats.

5 STATISTICAL METHODS

An outline of the logical flow of statistical methods used for this investigation is provided as a flow-chart in Fig. 2, for reference.

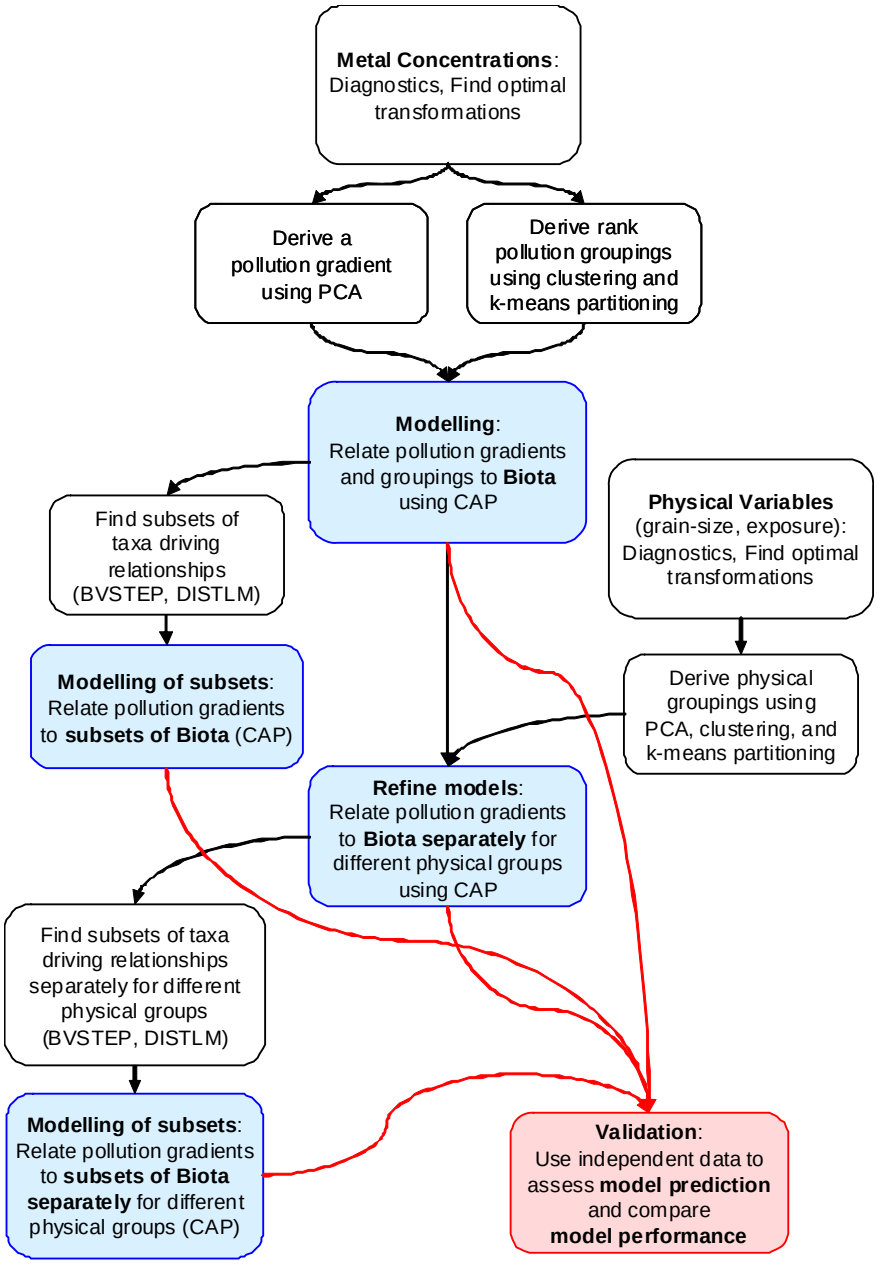


Fig. 2. Flowchart showing an outline of the logical flow of statistical analyses for the modelling (in blue) and validation (in red) procedures used in this investigation.

5.1 Pollution gradient based on metal concentrations

The individual distributions of concentrations of metals (Cu, Pb and Zn) in the 81 samples to be used for modeling were investigated using the R computer program (R Development Core Team 2005). These distributions were found to be strongly right-skewed, for both the <63 μm fraction and whole sample (<500 μm) values (see Appendix 4). Thus, before further analysis, an optimal transformation on the basis of the Box-Cox power series (Box and Cox 1964, Venables and Ripley 2002) was sought for each variable, which would render the data as close to normally distributed as possible to facilitate modeling. The Box-Cox power series transforms a variable y to $y' = (y^p - 1)/p$ for a given power p to obtain a more symmetric and normal distribution. For $p = 0$, the transformation is $y' = \ln(y)$. These equations rely on the values of y to be strictly positive, i.e. $y > 0$. More generally, one can use $y' = ((y + \lambda)^p - 1)/p$ (and $y' = \ln(y + \lambda)$ in the case of $p = 0$), for situations where $y > -\lambda$. In the present investigation, to allow for zero values, we used $\lambda = 1$. These power transformations are monotonic, which means they do not change the rank-order of the data points. To find an optimal transformation for a given variable, the log-likelihood of the normal distribution for the data is calculated for a series of values of p, and an optimal value for p is chosen to achieve normality where the maximum is achieved (i.e. the maximum likelihood). In the present case, optimal values for the power transformation ranged between 0.12 and 0.37; however, the natural log transformation ($y' = \ln(y)$) also resulted in more symmetrical unimodal distributions of the metal concentration variables in all cases (Appendix 4). Furthermore, principal component analyses done using logged data versus optimally power-transformed data yielded extremely similar results. Therefore, for simplicity in interpretation and also to retain the ability to back-transform results to original values with relative ease, log metal concentrations were used in all subsequent analyses.

There were very strong correlations among the three log metal concentration variables for both datasets. For the <63μm dataset, the correlations were: $r_{Cu,Pb}$ = 0.90, $r_{Cu,Zn}$ = 0.93 and $r_{Pb,Zn}$ = 0.95; for the <500μm dataset, the correlations were: : $r_{Cu,Pb}$ = 0.88, $r_{Cu,Zn}$ = 0.91 and $r_{Pb,Zn}$ = 0.94. It was therefore logical, in each case, to seek a single variable which would characterise an overall pollution gradient corresponding to increases in the concentrations of all three metals in the field. Thus, principal component analyses (PCAs) were done for each of the <63μm and <500μm datasets (81 samples) on the basis of log-transformed metal concentrations, using the PRIMER v6 computer program (Clarke and Gorley 2006). Due to the fact that the three log metal concentration variables (in each case) were on the same measurement scale and in the same units, PCAs were done on raw, rather than normalized data. Analyses based on the logged data versus on normalized logged data made no practical difference to the results. By doing the analysis on raw (logged) variables, we retained the relative ease of placing a

new (validation) object into the PCA space on the basis of the original metal concentrations measured at a site.

5.2 Analyses to identify pollution groupings

Rank-order groupings along any identified pollution gradient would be a useful tool for classifying sites in the context of environmental management. Three methods were used to identify possible groupings. Hierarchical agglomerative group-average clustering was done in PRIMER to obtain a dendrogram of samples on the basis of Euclidean distances using the three log metal concentration variables. Hierarchical clustering begins with each individual sample being separate. The two samples having the smallest Euclidean distance are joined together. The algorithm then progressively joins individual samples and, subsequently, groups of samples, in a hierarchical fashion, using the criterion of the minimum group-average distance. Once the dendrogram was complete, groupings were then identified using two methods: by taking an arbitrary slice through the dendrogram at a given distance value or by using similarity profiles (SIMPROF, e.g., Potter et al. 2001), as provided in PRIMER. SIMPROF calculates rank profiles of similarities among samples and determines the statistical significance of each individual split in a dendrogram on this basis using a permutation technique. The shape of the similarity profile under a true null hypothesis of no inherent structure among samples that would warrant a split is generated by permuting the values for each variable independently across all samples involved in that split. The essential concept here is that correlation structure among variables generates inherent structure among samples. For more details, see Potter et al. (2001) and Clarke and Gorley (2006). As for the PCAs above, all analyses were done separately for the <63 μm and <500 μm datasets.

In addition to the "slice" and SIMPROF methods of identifying groups, the third method used here was *k*-means partitioning (MacQueen 1967). This was done using a special-purpose FORTRAN program (courtesy of Pierre Legendre, University of Montreal). This method begins with all samples together in a single large group. For a given number of groups (*k*), it divides the samples into *k* groups so as to minimize the sum of squared distances of the samples to their group centroid (defined as the average for the variables within that group). The question then becomes: what value of *k* is optimal for a given dataset? The optimal number of groups was selected using the Calinski-Harabasz criterion (Calinski and Harabasz 1974). This criterion is defined for a given number of groups (*k*) as:

$$CH_k = \frac{\{R^2/(k-1)\}}{\{(1-R^2)/(N-k)\}}$$

where R^2 is the explained sum of squares, *N* is the total number of samples and *k* is the number of groups. For a given set of groups, the explained sum of squares will simply increase with increases in the number of groups. The CH_k criterion is therefore

standardised (essentially like an *F*-statistic), in order to take this into account. An appropriate number of groups (*k*) is chosen where this criterion is maximised. *K*-means partitioning solutions and associated CH_k values were calculated for each dataset (<63 μm and <500 μm) on the basis of Euclidean distances of log metal concentrations for each of $k = 2, 3, \ldots, 12$ groups.

The relevance of the groups obtained using each of the above three procedures was ascertained by visual examination on the PCA plots and also by calculating the value of each of two statistics designed to measure the degree of separation among groups: the ANOSIM *R*-statistic (Clarke 1993) and the PERMANOVA *F*-statistic (Anderson 2001). The *R*-statistic is a standardised test-statistic which varies from –1 to +1 and measures the difference in the average rank within-group similarities from the average rank between-group similarities. The larger the value (i.e., the closer the value is to +1) the more distinct are the groups. The pseudo *F*-statistic also provides a useful measure of group differences. When Euclidean distances are used, as for the metal data here, it is effectively the sum of the between-group sum of squares divided by the sum of the within-group sum of squares, each divided by appropriate degrees of freedom. Notably, in this case the PERMANOVA *F*-statistic is the same as the Calinski-Harabasz criterion described above. Although not scaled (its distribution does depend on the degrees of freedom), larger values of *F* indicate greater separation of groups. The primary difference between these two test-statistics is that the ANOSIM *R*-statistic is based on ranks of similarities (or distances), whereas the PERMANOVA *F*-statistic is based on the actual distances themselves.

5.3 Predicting the pollution gradient on the basis of biotic assemblages

Having identified a pollution gradient on the basis of log metal concentrations in either the <63 μm fraction (PC1.63) or the total sample (< 500 μm, PC1.500, see results), the next step was to determine if there was a significant relationship between the biotic assemblages and these gradients. In essence, we wished to know whether positions along the pollution gradient axes (either PC1.63 or PC1.500) could be obtained using the biotic data alone. This was done using canonical analysis of principal coordinates (CAP, Anderson and Robinson 2003, Anderson and Willis 2003). In general, traditional canonical analysis can be used to find a linear combination of variables that have maximum correlation with either (i) one or more continuous variables (gradient analysis) or (ii) a set of groups (discriminant analysis). More particularly, if we consider the biotic data as a multivariate cloud of sample points (where abundances of each taxon correspond to a separate dimension), can we find an axis through this cloud that (i) is most highly correlated with the pollution gradient or (ii) is best at separating rank pollution groupings? Such axes are called *canonical axes*.

We do not expect that the abundances of individual species or taxa (biotic variables) will necessarily have a direct linear relationship with log metal concentrations or with PCA-derived pollution axes. However, it is possible that log abundances of taxa could be

linearly related to such environmental gradients (e.g., Warton 2005). It is also possible that the relationship may be approximately linear through the use of an appropriate ecological dissimilarity measure. The CAP technique can be done on the basis of any dissimilarity measure of choice, as principal coordinate (PCO) axes (Gower 1966) are used instead of the original variables for the analysis. If CAP is done on the basis of Euclidean distance, then it is the same as a traditional canonical correlation analysis, seeking linear relationships between the two sets of variables (e.g., Legendre and Legendre 1998). However, if some other dissimilarity measure is used, then the distances among points in the multivariate cloud being examined is defined by the dissimilarity measure chosen. This means the relationship between the original taxa variables and the canonical axes will not be linear, but will generally be some complex unknown relationship that incorporates composition and abundance information, depending on how the dissimilarity measure is defined.

Canonical axes may not travel through the multivariate cloud in the same direction as the direction of greatest total variation, so some measure of their success at identifying and predicting real and repeatable patterns in the data is required, using validation techniques. One approach is to take out one sample at a time and apply the canonical model from all of the other samples to the "left-out" sample in order to place it into the canonical space and allocate it to a particular group or gradient position (e.g., Lachenbruch and Mickey 1968). The proportion of samples that were correctly allocated is called the *leave-one-out allocation success*. This can be compared with what would be expected from random allocation to assess the utility of the model. For example, with 2 groups, we would expect random allocation to provide an allocation success rate of ~50%. Similar measures of allocation success can be used to measure the utility of the model with new, or "validation" sites (see section 3.6 below). If the model being developed does not involve groups, but rather concerns a gradient, then the *leave-one-out residual sum of squares* performs a similar function, which we aim to minimize.

CAP was done to relate the biotic assemblages: (i) to each of the PC pollution gradient axes from the log metal concentration data (<63 μm and <500 μm) and (ii) to the 5 rank pollution groupings identified from *k*-means partitioning of the log metal concentration data (see results). The flexibility of the CAP approach allows a plethora of different possible approaches for modeling, depending on the dissimilarity measure used. Each approach will emphasize different aspects of the biotic community information available. With this in mind, four different dissimilarity measures were chosen and used as the basis of analyses:

- *Bray-Curtis dissimilarity on square-root transformed data*. This approach is generally highly recommended for the analysis of species abundance data in ecology (Clarke and Gorley 2006). It emphasizes both composition and relative abundance information, ignores joint absences and the modest transformation to square roots reduces the relative influence of highly abundant taxa.

- *Jaccard dissimilarity on presence/absence data*. This dissimilarity measure is interpretable as the percentage of unshared species (e.g., Chao et al. 2005) and thus addresses compositional information only.
- *Euclidean distance on ln(x+1)-transformed abundances*. A limitation of the Bray-Curtis measure is its lack of discrimination at its upper limit and its sporadic behaviour with sparse data (Clarke et al. 2006). The Euclidean distance does not have an upper limit and also does not ignore joint-absence information, thus allowing samples having few species to cluster together. This may be appropriate in the present study, where high levels of contaminants could generate sparse sample units. The log-transformation allows for the well-known fact that most species abundances have right-skewed distributions (e.g. Warton 2005).
- *Modified Gower dissimilarity*. This measure was only recently described (Anderson et al. 2006) and is directly interpretable as the average order-of-magnitude change in abundance on a log scale per taxon. It excludes joint absences but has no upper limit and, unlike Bray-Curtis, it therefore does not lose discrimination with decreasing overlap of shared species. In addition, a unit-change in composition (from 0-1) is treated with the same weight as an order-of-magnitude change in abundance. In the present study, a log to the base 10 was used to define an "order-of-magnitude".

In each case, the number of PCO axes used for the CAP analyses (*m*) was chosen so as to maximize the leave-one-out allocation success (in the case of discriminant analyses) or so as to minimize the leave-one-out residual sum of squares (in the case of gradient analyses).

5.4 Subsets of taxa that can be used for prediction

It was of interest to determine which species might be driving any relationship that might be found between the biotic assemblages and the pollution gradients (PC1.500 and PC1.63). There were two reasons for addressing this issue: (i) it is of biological interest to consider which taxa may be most sensitive to pollution gradients and (ii) a more efficient predictive model may be used for monitoring and management if a smaller number of taxa can serve as indicators of the pollution gradient.

The first approach was to consider a subset of taxa that was nominated by J. Hewitt (NIWA) *a priori* as being sensitive to pollution on the basis of known biological qualities regarding field distributions (in the Auckland Region) and laboratory toxicity studies. This consisted of 22 taxa, which are hereafter referred to as the *sensitivity* subset (more details are provided in the results). An *ecological* subset of 43 taxa (listed in Appendix 5), as had been nominated and used in the previous study by Anderson et al. (2002) and Hewitt et al. (2005), was also considered.

The next approach was simply to examine scatterplots of individual taxa versus each of PC1.500 and PC1.63 across the sites and to choose those taxa which appeared to be

responding in some fashion (either positively or negatively) on the basis of this visual inspection. Although this may be thought of as a kind of "data snooping" (and it is!), we are not in the business of trying to then "test the significance" of any resulting relationship that might be obtained on the basis of such a subset (an illogical procedure). Rather, the subset, once obtained was examined for its ability to predict the positions of sites along the gradient, which is an entirely different issue.

A third approach was to use the BVSTEP routine provided in PRIMER. This routine seeks to find subsets of variables which will produce a dissimilarity matrix having the strongest rank correlation (Spearman's rho, ρ) with a fixed distance matrix representing some particular model of interest. In our case, the model of interest is Euclidean distances among samples based on either PC1.500 or PC1.63. The BVSTEP routine begins by randomly selecting a set of, say, 6 variables. It then gradually adds and/or drops variables in a step-wise fashion so as to maximize the correlation between the dissimilarity matrix obtained from those variables and the Euclidean distance matrix from the pollution gradient. Although it would be ideal to search over all possible combinations of variables, this is prohibitively time consuming for anything beyond about 17 variables, so it was out of the question to attempt it for 102 taxa. However, BVSTEP may find solutions that are only local maxima, as opposed to a global maximum value for ρ. Thus, in each case, we ran BVSTEP on the basis of 6 initial random variables and 20 random starts to increase the chances of finding true global maximum values. The BVSTEP routine was run on the basis of three dissimilarity measures: (i) Bray-Curtis, after square-root transformation, (ii) Euclidean distance after ln(x+1) transformation and (iii) Modified Gower (log base 10) and was run separately for PC1.63 and PC1.500.

Finally, another potential optimal subset for modeling was obtained by performing a step-wise selection of ln(x+1)-transformed taxa in a linear model to predict Euclidean distances among points along the pollution gradient, using DISTLM (McArdle and Anderson 2001). This was deemed reasonable on the basis that the CAP analyses using Euclidean distance on ln(x+1)-transformed abundances performed as well or better than any other approach (see results). The selection criterion used at each step was the Bayesian Information Criterion (BIC), proposed by Schwarz (1978). This measure balances the value of the log-likelihood with a penalty for the number of parameters used in the model (e.g., Seber and Lee 2003). We used this criterion, rather than Akaike's "An Information Criterion" (AIC, Akaike 1973), because the AIC is known to have a tendency to overfit (e.g., Seber and Lee 2003). Smaller BIC values indicate a better model fit. This criterion is defined in the present case as:

$$BIC = N \times \ln(SS_{RES} / N) + (\ln(N) \times p)$$

where N is the total number of samples, p is the number of parameters or variables used in the linear model and SS_{RES} is the residual sum of squares. Step-wise selection proceeds iteratively by first a forward selection and then an attempted backwards elimination of variables so as to enhance (reduce) the value of the BIC criterion. The procedure stops when no additions or deletions of variables will decrease the BIC. This

analysis was done using a beta version of the PERMANOVA+ add-on package to the PRIMER computer program (to be released in 2007).

After all of the proposed subsets had been determined, their performance was compared by calculating: (i) the value of the Spearman rank correlation (ρ) between the dissimilarity matrix obtained using the subset of taxa and the Euclidean distance matrix of each pollution gradient (calculated using the routine called RELATE in the PRIMER program); and (ii) a sequential DISTLM, regressing each of the pollution gradients versus each of the subsets, including determination of the percentage of the variation explained (Note: the latter was only done for those subsets which were derived using a Euclidean distance approach).

Where the relationship using a particular subset was deemed to be substantial enough (i.e., rivaling the results obtained on the basis of all taxa), a CAP analysis was done, including relevant diagnostics, to predict the pollution gradient on the basis of the subset.

5.5 Refinements using grain-size characteristics and exposure indices

Results indicated that the sites identified as most polluted were generally located in the upper reaches of estuaries in some of the least exposed locations (see results). In addition, the sensitivities of organisms characterising sites that have different sediment textures and exposures may vary considerably. The model developed by direct measures of metal concentrations along a gradient should not be entirely confounded by a concomitant gradient in sediment texture and/or exposure. The available information regarding physical sediment characteristics and exposure at the sites should be used to inform and enhance models as much as possible.

We considered two possible ways to incorporate sediment texture and exposure variables into the modeling process. The first was to treat the physical characteristics as covariates. That is, essentially to "remove" the known relationships (statistically) between the biota and the physical variables before investigating the relationship between the biota and the pollutants on the information that remains (residuals). The second was to seek a partitioning of the dataset into groups on the basis of the physical variables and to then examine relationships between the biota and the pollutants separately within each of those groups.

First, diagnostics of the physical variables (the 4 grain-size variables and the 2 exposure variables) were done and an optimal transformation to normality for each was sought using the Box-Cox transformation, as had been done for the metal concentration variables (Box and Cox 1964, see section 3.1 above for more details). The correlations among all of the physical variables and PC1.500 and PC1.63 were also examined in detail. Not surprisingly, a strong correlation was revealed between the percentage of silt and clay (the mud fraction, < 63 μm) and PC1.500 ($r = 0.73$). This meant that the proposed approach of "removing" the relationship between the biota and sediment

texture would also effectively remove most of the information inherent in the metal concentration data itself, tantamount to "throwing the baby out with the bathwater", so to speak. Furthermore, there is no currently known statistical method that can achieve a "removal" of effects of physical variables when biotic variables are treated as *predictors* in an indirect non-linear fashion through a chosen dissimilarity measure, as in the present case[1]. Thus, the approach of partitioning the dataset into groups based on the physical variables was chosen as a more appropriate route, especially in keeping with the ultimate goal of being able to make successful predictions for new sites in these systems.

A principal component analysis was done to visualise the relationships among sites in terms of the six optimally transformed physical variables. Data were first normalized (by subtracting the mean and dividing by the standard deviation) because, unlike the metal concentration variables, these variables did not have comparable scales of variation. A search for appropriate groupings of sites was done using three techniques, as had been done for the metal concentration data: (i) hierarchical agglomerative group-average clustering and the use of SIMPROF; (ii) an arbitrary slice through the resulting dendrogram and (iii) *k*-means partitioning, using the Calinski-Harabasz criterion to identify an appropriate number of groups. The resulting groupings were then superimposed on the PCA plot and were also subjected to analysis by ANOSIM and PERMANOVA to determine the most parsimonious grouping of sites on the basis of the physical variables.

Two groups of sites were identified using the above procedure (see the results for more details). For each of these two groups (one corresponded to sites with generally finer sediments and one with generally coarser sediments), the relationship between the biota and each of PC1.500 and PC1.63 was investigated using CAP on the basis of each of the following measures: (i) Bray-Curtis dissimilarities on square-root transformed abundances, (ii) Euclidean distances on ln(x+1)-transformed abundances and (iii) Modified Gower dissimilarities with log base 10. In each case, the canonical correlation and leave-one-out residual sum of squares were investigated to identify the best of these models. Furthermore, the BVSTEP and DISTLM procedures were used, as previously described, to identify subsets of species that could be driving observed relationships. Subsequent CAP analyses were also done on these subsets and results compared.

5.6 3.6 Validation

The biological, chemical and physical data were set aside for 14 samples from the modeling process so that these could be used for model validation (see Fig. 1, Appendix 1). We wished to determine how closely the best models were able to place

[1] dbRDA *can* be used to achieve such a removal, however, when biotic variables based on a dissimilarity measure are treated as the *responses* rather than as the *predictors*.

each new validation sample onto an existing canonical axis and, from this, to predict the true position of that sample along the relevant pollution axis.

First, the chemical data were used to place each sample point onto each of the pollution axes (PC1.500 and also PC1.63, as obtained using PCA, see results) using the actual metal concentrations measured at each site. These positions were deemed to be the true or "actual" values for those sites along each pollution axis. Models which showed strong canonical correlations as well as small leave-one-out residual sums of squares were then each used, in turn, to place each validation point onto the pollution axes. For each model, these were the "predicted" values along the pollution gradient. The sum of squared deviations of the predicted values from the actual values (the residual sum of squares, SS_{RES}) was then calculated for each model. The models with the smallest values for SS_{RES} were considered to have achieved the best predicted fit. Scatter plots of the predicted versus the actual values were also used to identify the sites whose predicted values deviated the most from their actual values and in which direction.

Models which had been refined by including physical data required allocation of the validation sites to one of the two physical groupings that had been identified (see results for details). This was done visually by superimposing the validation sites onto the principal component plot based on the physical data. Identical results were obtained using a canonical analysis approach to the allocation (results not shown). After allocating validation sites to a physical grouping, separate canonical analyses were done accordingly and SS_{RES} was calculated, as for the other models.

6 RESULTS

6.1 Identification of a pollution gradient using PCA

The first principal component explained approximately 94% of the variation in log metal concentrations for the <500 μm dataset and just over 95% of the variation in log metal concentrations for the <63 μm dataset (Fig. 3). The value of a particular site at a particular time along the first PC axis, in each case, therefore provides an excellent measure of the degree of pollution (amount of metals) in the sediment at that site and time. We shall refer to the first principal component for these two datasets as PC1.500 and PC1.63 in what follows. The sign of a principal component axis is arbitrary, so the axis was deliberately orientated such that increasing values along PC1, in each case, would correspond to increasing metal concentrations, for convenience in interpretation (Fig. 3).

For a principal component analysis, the eigenvector weights provide coefficients for a linear combination of the (centered) original variables that will yield the principal component scores. For the PC axes shown in Fig. 3, we obtained:

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

$$\text{PC1.63} = 0.524\times(X_{Cu}^{(63)}) + 0.534\times(X_{Zn}^{(63)}) + 0.663\times(X_{Pb}^{(63)})$$

where, for example, $X_{Cu}^{(500)}$ = the log concentration of copper in the total sample (<500 μm) minus the mean log concentration of copper (<500 μm) across the full set of 81 samples, and so on for the other variables. The means that should be used to obtain centered values for each of the variables are shown in Table 1.

Table 1. Mean log concentrations for each metal in each dataset for the 81 samples used for modeling.

Metal	Mean (<500 μm)	Mean (<63 μm)
Cu	2.472	2.876
Zn	4.418	4.643
Pb	2.925	3.327

Thus, for example, for a given validation site, if the metal concentration values from the <500 μm fraction are given as $\{x_{Cu}, x_{Zn}, x_{Pb}\}$ for copper, zinc and lead, respectively, then the value for that site along the pollution gradient axis would be:

$$\text{PC1.500} = 0.615\times(\ln(x_{Cu}) - 2.472) + 0.528\times(\ln(x_{Zn}) - 4.418) + 0.586\times(\ln(x_{Pb}) - 2.925)$$

The ordered values of particular samples (sites within a particular year) along pollution gradient axis PC1.500 are shown in Table 2. Similar information for PC1.63 is shown in Table 3.

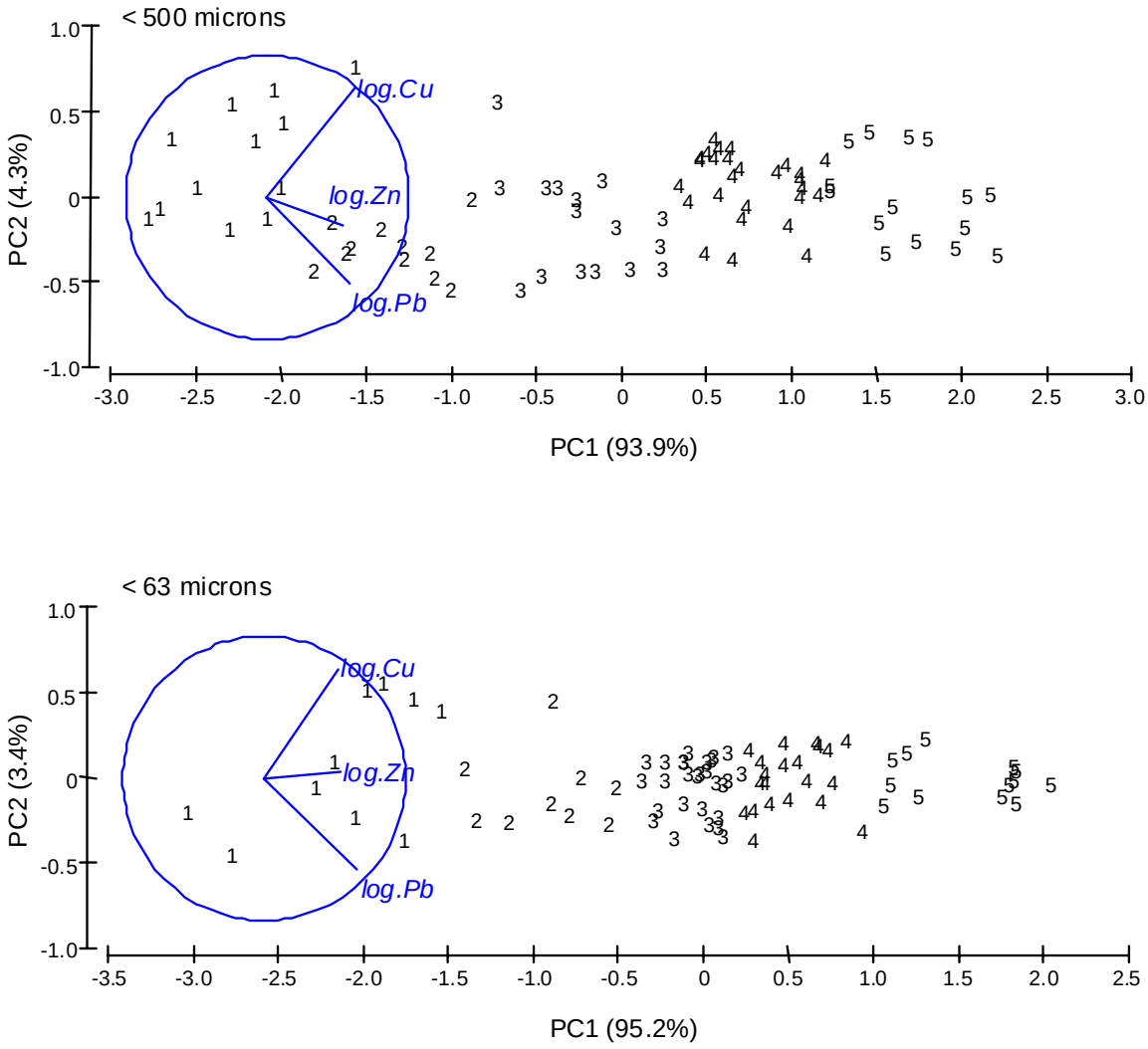


Fig. 3. PCA of log metal concentrations (Cu, Pb and Zn) for the whole sample (< 500 μm, top) and for the mud fraction (<63 μm, bottom) at each of 81 sites. Numbers indicate the 5 groups (1 = healthy, 5 = polluted) obtained using k-means partitioning on the basis of Euclidean distances.

Table 2. List of the samples used for modeling, ordered on the basis of increasing values along the pollution gradient represented by principal component axis 1 from the whole sediment sample (PC1.500). Also given is the group identification of each sample based on the 5-group *k*-means partitioning.

Year	Site no.	Site name	PC1.500	Group
2006	47	Okura D	-2.7806	1
2002	9	Clarkes Beach	-2.7159	1
2004	50	Orewa G	-2.6402	1
2002	7	Cape Horn	-2.4859	1
2005	23	Hobsonville	-2.3006	1
2004	62	Puhoi F	-2.2836	1
2004	49	Orewa F	-2.1429	1
2004	31	Mangemangeroa B	-2.0854	1
2004	75	Waiwera J	-2.0476	1
2005	12	Glendowie	-2.0120	1
2004	74	Waiwera E	-1.9847	1
2004	38	Meola Outer	-1.8100	2
2002	23	Hobsonville	-1.7070	2
2005	27	Lower Shoal Bay	-1.6113	2
2004	70	Turanga G	-1.5901	2
2004	63	Puhoi H	-1.5711	1
2002	61	Puhinui, Entrance	-1.4129	2
2004	69	Shoal Bay, Upper	-1.2919	2
2004	73	Victoria Ave	-1.2734	2
2004	8	Chelsea	-1.1211	2
2005	22	Hobson - Tohunga	-1.0933	2
2004	10	Coxes, Waitemata	-0.9955	2
2005	52	Out Main UWH	-0.8843	2
2005	76	Weiti	-0.7310	3
2005	60	Puhinui	-0.7136	3
2005	43	Motions East	-0.5960	3
2004	15	Henderson Entrance	-0.4685	3
2004	20	Hillsborough	-0.4386	3
2005	19	Hi North	-0.3763	3
2004	4	Bengazi	-0.2657	3
2005	64	Pukaki	-0.2603	3
2002	39	Meola Reef	-0.2455	3
2002	15	Henderson Entrance	-0.1683	3
2004	71	Turanga J	-0.1141	3
2005	58	Pollen Island	-0.0399	3
2005	77	Whakataka	0.0490	3
2005	39	Meola Reef	0.2281	3
2005	28	Lucus outer	0.2360	3
2002	77	Whakataka	0.2371	3
2005	14	Hellyers outer	0.3413	4
2005	13	Hellyers	0.3992	4
2005	72	Upper main UWH	0.4544	4
2005	35	Mangere Inlet: Kiwi Esplanade	0.4583	4
2004	3	Awatea Rd	0.4820	4
2005	56	Paremoremo	0.5013	4
2005	67	Rangitopuni UWH	0.5280	4

Year	Site no.	Site name	PC1.500	Group
2005	30	Lucus Upper	0.5386	4
2004	79	Whau Entrance	0.5675	4
2005	66	Rangitopuni	0.5742	4
2005	57	Paremoremo upper	0.6224	4
2005	6	Brigham	0.6372	4
2004	29	Lucus Te Wharau	0.6478	4
2005	40	Meola West	0.6540	4
2005	34	Mangere Inlet: Harania Creek	0.6903	4
2004	68	Shoal Bay, Hillcrest	0.7121	4
2004	59	Princess St	0.7333	4
2005	54	Pakuranga mid	0.9105	4
2005	33	Mangere Cemetery	0.9670	4
2005	45	Ngataringa Bay	0.9830	4
2002	41	Middlemore	1.0451	4
2004	16	Henderson Lower	1.0492	4
2004	55	Panmure	1.0547	4
2005	24	Kaipatiki	1.0610	4
2004	65	Purewa	1.0888	4
2004	51	Otahuhu Creek	1.1593	4
2005	1	Anns Creek	1.1863	4
2004	5	Bowden Rd	1.2195	5
2005	41	Middlemore	1.2207	5
2002	17	Henderson Upper	1.3310	5
2002	1	Anns Creek	1.4536	5
2005	78	Whau East	1.5129	5
2005	37	Meola Inner	1.5486	5
2005	46	Oakley	1.5911	5
2004	82	Whau Upper	1.6878	5
2002	42	Motions	1.7307	5
2005	84	Whau West	1.8031	5
2002	37	Meola Inner	1.9672	5
2005	82	Whau Upper	2.0153	5
2005	83	Whau Wairau	2.0243	5
2002	83	Whau Wairau	2.1583	5
2005	42	Motions	2.1976	5

Table 3. List of the samples used for modeling, ordered on the basis of increasing values along the pollution gradient represented by principal component axis 1 from the < 63 micron fraction (PC1.63). Also given is the group identification of each sample based on the 5-group k-means partitioning.

Year	Site no.	Site name	PC1.63	Group
2002	7	Cape Horn	-3.0239	1
2005	60	Puhinui	-2.7719	1
2004	50	Orewa G	-2.2785	1
2004	49	Orewa F	-2.1674	1
2002	9	Clarkes Beach	-2.0425	1
2004	62	Puhoi F	-1.9768	1
2004	63	Puhoi H	-1.8811	1
2002	61	Puhinui, Entrance	-1.7555	1
2004	75	Waiwera J	-1.6992	1
2004	74	Waiwera E	-1.5414	1
2004	47	Okura D	-1.3977	2
2005	64	Pukaki	-1.3297	2
2004	31	Mangemangeroa B	-1.1439	2
2004	70	Turanga G	-0.9028	2
2005	76	Weiti	-0.8860	2
2004	71	Turanga J	-0.7945	2
2005	35	Mangere Inlet: Kiwi Esplanade	-0.7189	2
2004	8	Chelsea	-0.5538	2
2004	20	Hillsborough	-0.5121	2
2005	34	Mangere Inlet: Harania Creek	-0.3598	3
2005	66	Rangitopuni	-0.3386	3
2005	58	Pollen Island	-0.2881	3
2005	27	Lower Shoal Bay	-0.2753	3
2005	56	Paremoremo	-0.2276	3
2002	23	Hobsonville	-0.2209	3
2004	73	Victoria Ave	-0.1703	3
2005	23	Hobsonville	-0.1263	3
2005	14	Hellyers outer	-0.1236	3
2005	12	Glendowie	-0.1168	3
2005	72	Upper main UWH	-0.0915	3
2005	67	Rangitopuni UWH	-0.0868	3
2004	79	Whau Entrance	-0.0379	3
2004	29	Lucus Te Wharau	-0.0267	3
2005	77	Whakataka	-0.0127	3
2005	33	Mangere Cemetery	0.0131	3
2005	52	Out Main UWH	0.0195	3
2004	68	Shoal Bay, Hillcrest	0.0311	3
2004	4	Bengazi	0.0399	3
2005	6	Brigham	0.0586	3
2004	69	Shoal Bay, Upper	0.0725	3
2005	45	Ngataringa Bay	0.0918	3
2002	77	Whakataka	0.0923	3
2005	22	Hobson - Tohunga	0.1102	3
2002	15	Henderson Entrance	0.1229	3
2004	15	Henderson Entrance	0.1402	3
2005	57	Paremoremo upper	0.1446	3

Year	Site no.	Site name	PC1.63	Group
2005	19	Hi North	0.2332	3
2004	65	Purewa	0.2411	4
2005	30	Lucus Upper	0.2683	4
2004	3	Awatea Rd	0.2867	4
2004	38	Meola Outer	0.2876	4
2004	16	Henderson Lower	0.3385	4
2005	24	Kaipatiki	0.3413	4
2005	43	Motions East	0.3585	4
2005	28	Lucus outer	0.3634	4
2005	39	Meola Reef	0.3833	4
2004	55	Panmure	0.4707	4
2002	1	Anns Creek	0.4707	4
2005	40	Meola West	0.4928	4
2004	51	Otahuhu Creek	0.5485	4
2005	13	Hellyers	0.6032	4
2004	5	Bowden Rd	0.6632	4
2002	17	Henderson Upper	0.6727	4
2005	84	Whau West	0.6908	4
2004	59	Princess St	0.7328	4
2002	39	Meola Reef	0.7631	4
2005	54	Pakuranga mid	0.8334	4
2004	10	Coxes, Waitemata	0.9336	4
2005	78	Whau East	1.0663	5
2005	46	Oakley	1.0892	5
2002	41	Middlemore	1.1132	5
2005	41	Middlemore	1.2050	5
2002	37	Meola Inner	1.2704	5
2005	1	Anns Creek	1.3117	5
2002	42	Motions	1.7486	5
2002	83	Whau Wairau	1.7893	5
2005	82	Whau Upper	1.8205	5
2005	37	Meola Inner	1.8318	5
2005	83	Whau Wairau	1.8355	5
2005	42	Motions	1.8410	5
2004	82	Whau Upper	2.0438	5

6.2 Groupings identified using clustering and *k*-means partitioning

Hierarchical agglomerative group-average cluster analysis on log metal concentrations from the whole sample (<500 μm) revealed some fairly clear groupings in the dendrogram (Fig. 4). An arbitrary slice at a Euclidean distance value of 1.05 yielded 5 groups. SIMPROF analysis indicated that there were 6 distinct groups (identified by the finest stems in the dendogram that remain in black, as opposed to those in red, Fig. 4). Analysis by *k*-means partitioning indicated that the most appropriate grouping structure was achieved by the partitioning solution found for $k = 5$ groups, which obtained the highest value of $CH_k = 191.77$ (Table 4).

The results obtained using metal concentrations for the <63 micron fraction also suggested some fairly clear groupings (Fig. 5). An arbitrary slice at a Euclidean distance value of 0.90 yielded 5 groups, although the number of samples per group was less evenly distributed than was seen for the whole sample analysis. The analysis using SIMPROF, however, suggested that there were 12 distinct groups (Fig. 5). The reason for this large number of groups could be due to the high correlation structure among the variables, which SIMPROF uses to characterize the presence of any structure worth distinguishing with further splits when assessing a given sub-set of samples. Although the best *k*-means solution in this case was found for $k = 3$ groups ($CH_k = 164.72$), comparable results were obtained for the 5-group partitioning solution ($CH_k = 163.69$, Table 4).

Table 4. Values of the Calinski-Harabasz criterion (CHk) for the k-means partitioningsolutions from k = 2-12 groups for each of the log metal concentration datasets.

No. groups (k)	<500 μm CH_k	<63 μm CH_k
2	184.34	120.02
3	176.33	164.72
4	189.20	157.76
5	191.77	163.69
6	169.52	160.88
7	142.94	162.41
8	164.25	146.80
9	150.09	156.32
10	137.21	125.13
11	124.22	112.00
12	114.66	97.03

The 3 different grouping solutions for each of the datasets were subjected to further analysis using the ANOSIM and PERMANOVA test-statistics (see Table 5). In the case

of the <500 μm dataset, the *k*-means 5-group solution obtained the highest value for both the ANOSIM and PERMANOVA test statistics. In addition, apart from just one of the samples (site 63, Puhoi H, 2004, see Table 2), the samples belonging to these five groups were ordered without any overlap along PC1.500 (Fig. 3). The results were a little less clear for the < 63 μm dataset: the grouping solution obtained using the "slice" method achieved the highest value of the *R*-statistic, whereas the *k*-means solution achieved the highest value of the PERMANOVA *F*-statistic (Table 5). However, the *k*-means 5-group solution also resulted in an ordering of samples without any overlap along PC1.63 (Table 3, Fig. 3). Thus, the 5-group *k*-means solution (with samples ranked from 1-5, with 1 = healthy and 5 = polluted) were used for subsequent analyses and plots, as shown in Tables 2 and 3 and in Fig. 3 for each of the datasets.

Table 5. Values of the ANOSIM R-statistic and the PERMANOVA F-statistic for each of the log metal concentration datasets and each of the grouping solutions.

	<500 μm			<63 μm		
			PERMANOVA			
	k	ANOSIM *R*	*F*	*k*	ANOSIM *R*	PERMANOVA *F*
k-means	5	0.910	191.77	5	0.859	163.69
Slice	5	0.907	175.64	5	0.913	139.54
SIMPROF	6	0.884	173.36	12	0.890	123.86

Maps which show each site colour-coded according to its pollution grouping according to the most recent sampling (from 1 = healthy to 5 = polluted) as indicated in Tables 2 and 3 are provided in Figures 6 and 7 for PC1.500 and PC1.63, respectively.

Having identified appropriate cluster groupings along the pollution gradient, we can therefore define classification boundaries along each PC gradient as being half-way between the highest value obtained along the axis for one group and the lowest value obtained along the axis for the next group. These proposed boundaries for classification are identified for each of the pollution gradient PC axes in Table 6. If a site obtains a value very far outside the overall minimum or maximum values (e.g., -3.5 or 3.0), then it would be considered outside the bounds of the current models for assessment purposes.

Table 6. Boundaries for classification along each of the pollution gradient PC axes.

		PC1.500		PC1.63	
	Group	min	max	min	max
healthy	1	-2.781	-1.897	-3.024	-1.470
↓	2	-1.897	-0.808	-1.470	-0.436
	3	-0.808	0.289	-0.436	0.237
	4	0.289	1.203	0.237	1.000
polluted	5	1.203	2.198	1.000	2.044

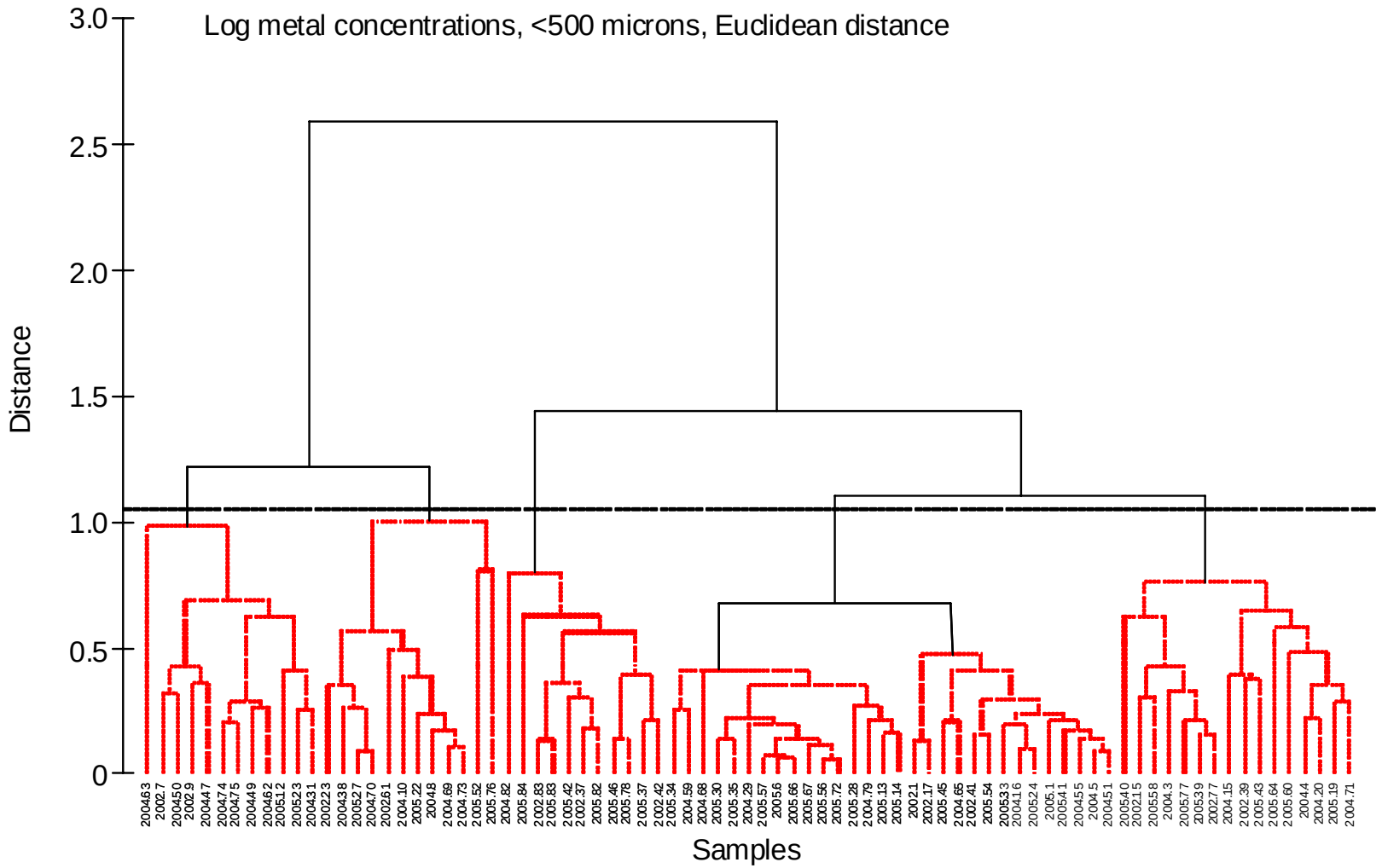


Fig. 4. Hierarchical agglomerative group-average cluster analysis among samples on the basis of log metal concentrations in the whole sample (< 500 microns) using Euclidean distances. The splits in red were not statistically significant, by SIMPROF, which indicated 6 groups. The single horizontal line indicates an arbitrary slice at a distance value of 1.05 which yielded 5 groups. Labels indicate the year and site number (as per Appendix 1) for each sampling unit.

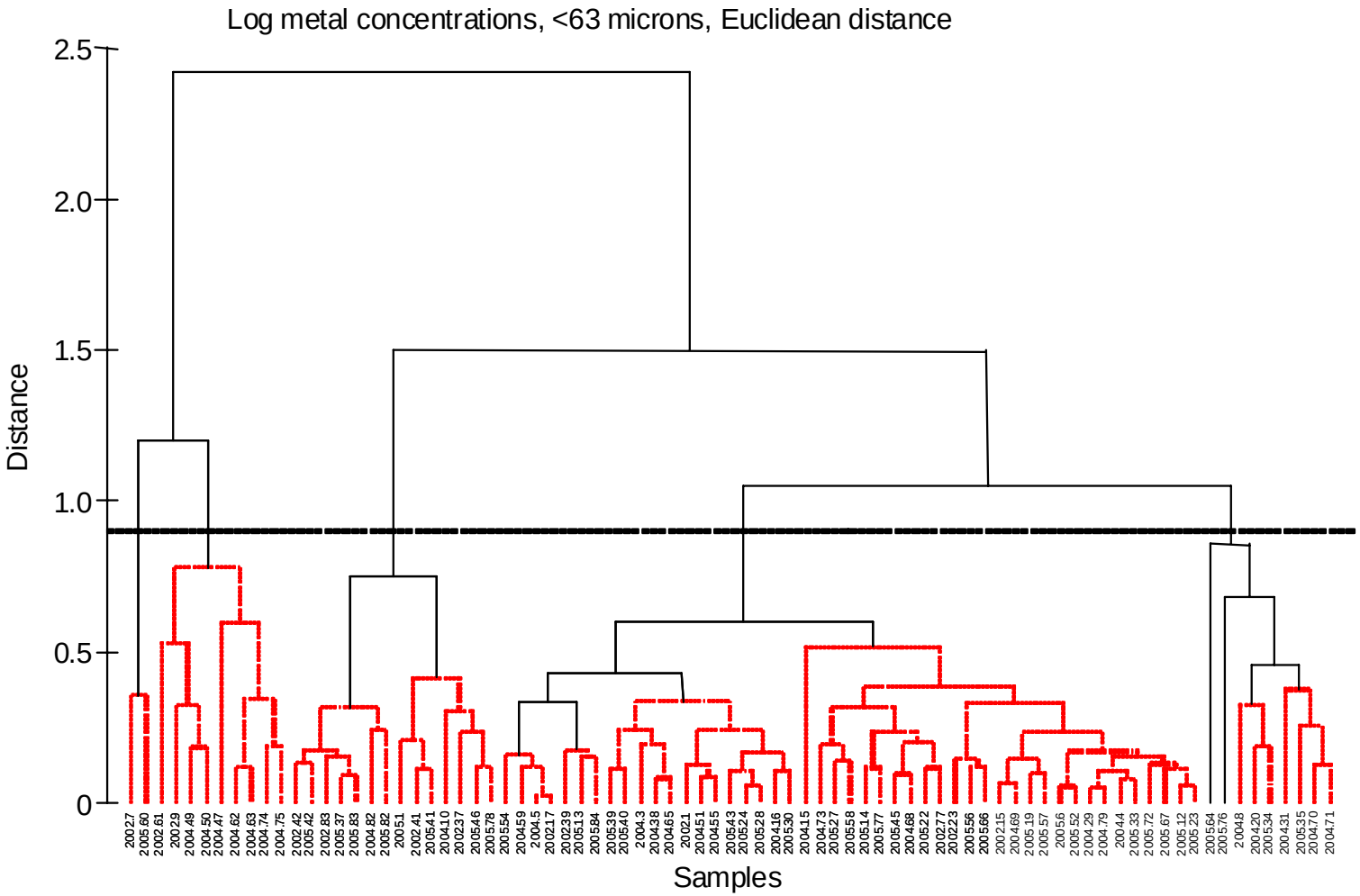


Fig. 5. Hierarchical agglomerative group-average cluster analysis among samples on the basis of log metal concentrations in the <63 micron fractio using Euclidean distances. The splits in red were not statistically significant, by SIMPROF, which indicated 12 groups of samples. The single horizontal line indicates an arbitrary slice at a distance value of 0.9 which yielded 5 groups. Labels indicate the year and site number (as per Appendix 1) for each sampling unit.

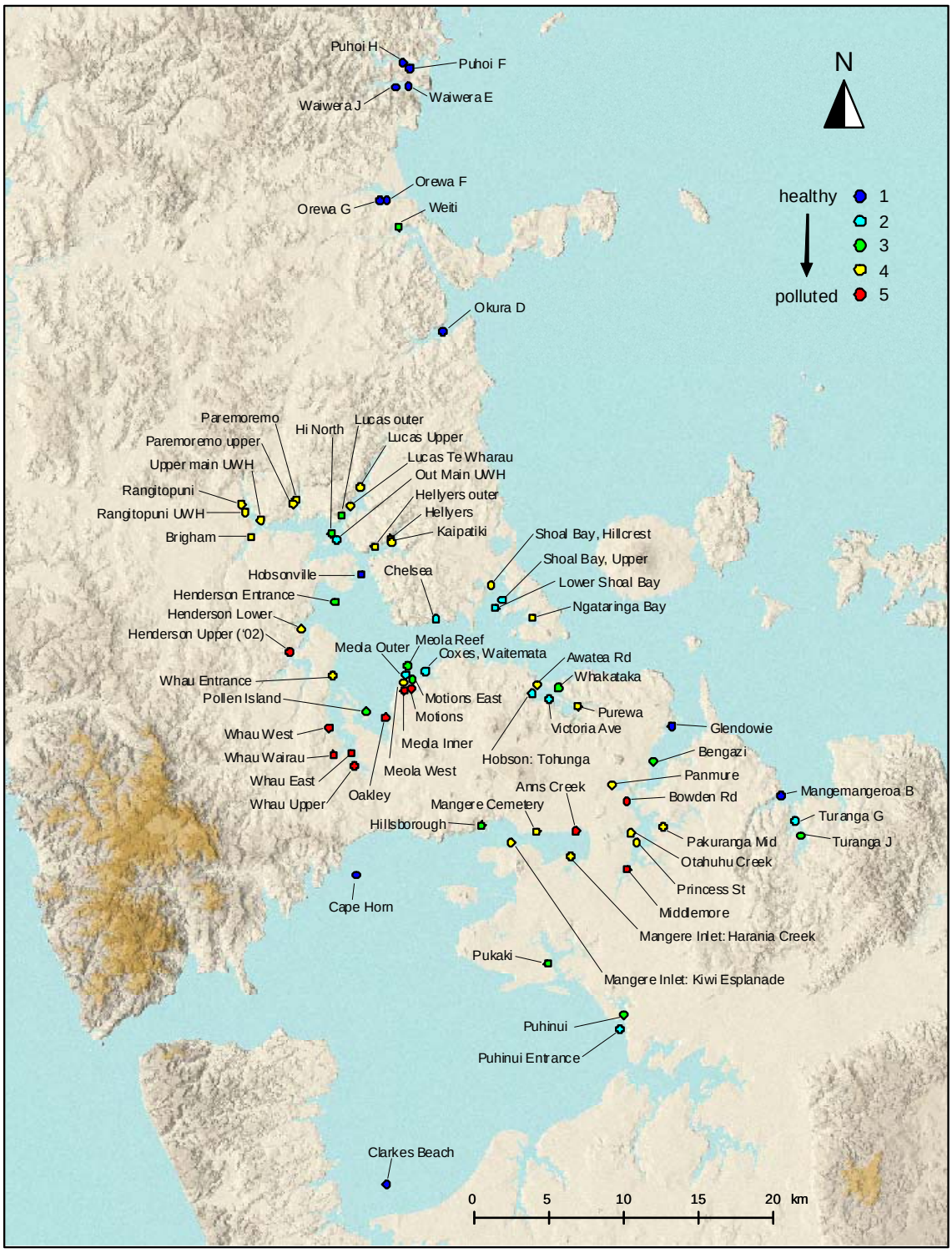


Fig. 6. Map of the Auckland region showing pollution groupings on the basis of *k*-means partitioning (1 = healthy, 5 = polluted) of the first principal component of log metal concentrations (Cu, Pb and Zn) for the whole sample (< 500 μm) at each of 81 sites.

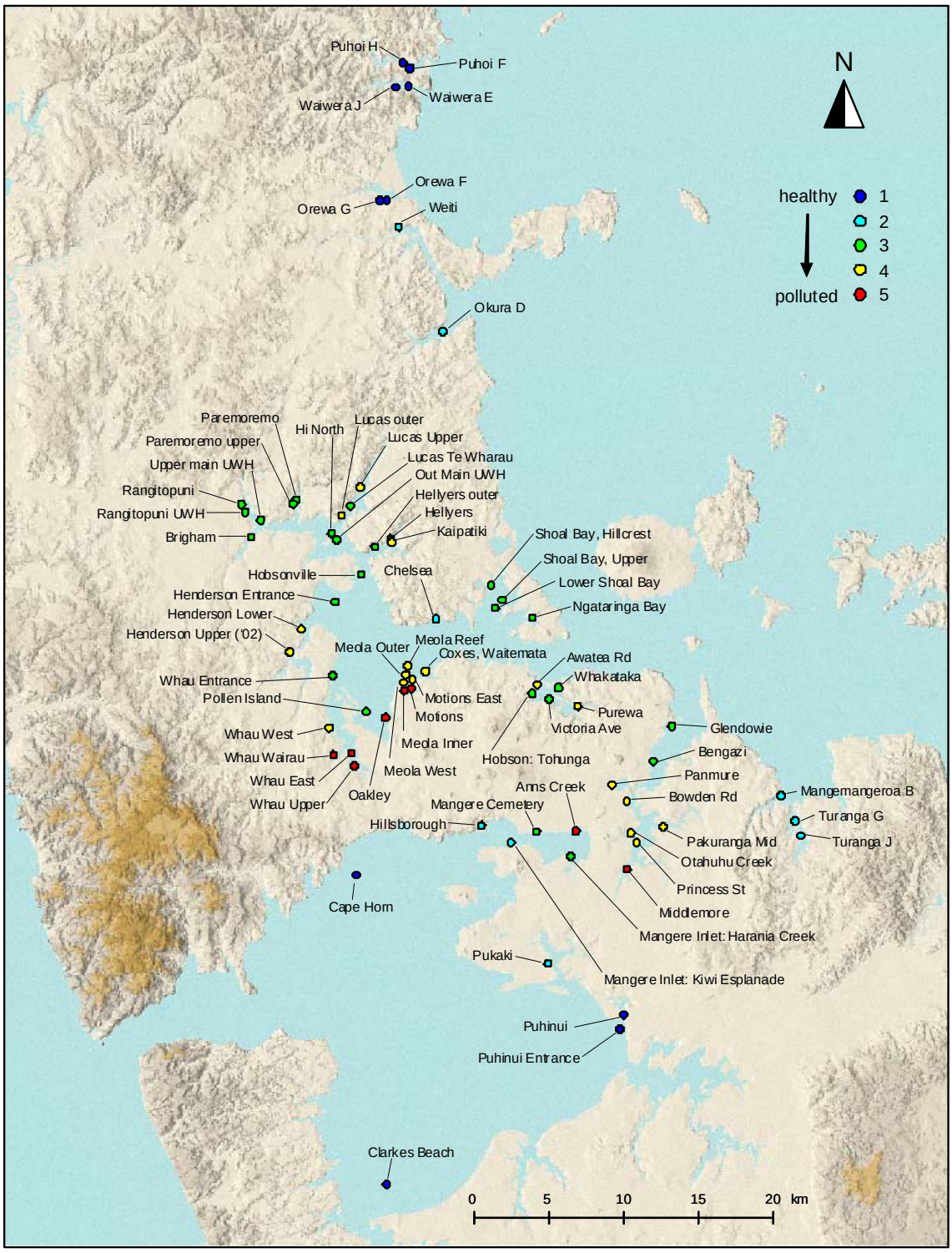


Fig. 7. Map of the Auckland region showing pollution groupings on the basis of *k*-means partitioning (1 = healthy, 5 = polluted) of the first principal component of log metal concentrations (Cu, Pb and Zn) for the mud fraction (< 63 µm) at each of 81 sites.

6.3 Predicting the pollution gradient on the basis of biotic assemblages

A summary of the CAP analyses designed to discriminate among the 5 groups for each of the <63 μm and <500 μm log metal concentration groupings and on the basis of each of the dissimilarity measures is shown in Table 7.

Table 7. Summary of CAP analyses to discriminate among the 5 rank pollution groupings (identified using *k*-means on log metal concentrations) using 102 taxa on the basis of each of four different dissimilarity measures, as indicated. Results are given separately for the metal concentrations from the mud fraction (<63 μm) and from the total sample (<500 μm). *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, SS_{RES} = the leave-one-out residual sum of squares, δ_i is the squared canonical correlation for the *i*th canonical axis and %correct = the leave-one-out allocation success of each model.

< 63 μm

	m	prop.G	SS_{RES}	δ_1	δ_2	δ_3	δ_4	%correct
BC, sqrt	11	0.836	3.481	0.633	0.437	0.268	0.076	54.32
Jaccard	12	0.762	3.599	0.519	0.467	0.236	0.070	51.85
Euc, ln(x+1)	10	0.826	3.723	0.614	0.318	0.160	0.082	46.91
Mod. Gower	27	0.931	3.707	0.745	0.670	0.394	0.320	56.79

< 500 μm

	m	prop.G	SS_{RES}	δ_1	δ_2	δ_3	δ_4	%correct
BC, sqrt	12	0.858	3.332	0.764	0.400	0.214	0.119	56.79
Jaccard	12	0.762	3.333	0.749	0.408	0.188	0.095	56.79
Euc, ln(x+1)	19	0.943	3.664	0.794	0.542	0.277	0.166	54.32
Mod. Gower	25	0.914	3.443	0.821	0.510	0.387	0.295	59.26

Unfortunately, none of these analyses were very convincing, as all had allocation success rates of < 60%, regardless of the dissimilarity measure used (Table 7). Although the percentages of correct allocation were considerably better than random allocation (which we would expect to achieve ~20% success for 5 groups), these models of the rank pollution groupings alone did not yield a very useful tool for predicting ecosystem health.

In contrast, canonical analyses which related the biota to the PC pollution gradient axes directly were much more successful (Table 8). The relationships between each of the PC pollution axes and the biotic data measured using any of the four chosen dissimilarity measures were all approximately linear, so it was not necessary to consider the more complex approach of using some form of nonlinear canonical analysis (e.g., NCAP Millar et al. 2005). Overall, it was clear that there was a stronger relationship between the faunal data and the pollution gradient based on the analysis of

metals from the whole sample (i.e. <500 μm) than there was between the faunal data and the pollution gradient based on metals obtained from the mud fraction alone (< 63 μm). In both cases, the relationship between the fauna and the PC axes was strongest using either Euclidean distances on ln(x+1)-transformed abundances or the Modified Gower measure (Table 8). However, the Modified Gower approach did require a large number of PCO axes to obtain such a good relationship. As predicted by general ecological theory, all of the dissimilarity measures that used relative abundance information performed better than the Jaccard measure (based on presence/absence) for these models.

Table 8. Summary of CAP analyses to model pollution gradients (obtained from PCA on log metal concentrations) using faunal data (102 taxa) on the basis of each of four different dissimilarity measures, as indicated. Results are given separately for the metal concentrations from the mud fraction (< 63 μm) and from the total sample (< 500 μm). *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, SS_{RES} = the leave-one-out residual sum of squares, δ_1 is the squared canonical correlation for the canonical axis, correl = the correlation between the canonical axis and the pollution gradient.

< 63 μm

	m	prop.G	SS_{RES}	δ_1	correl
BC, sqrt	12	0.858	0.536	0.619	0.787
Jaccard	9	0.668	0.698	0.455	0.675
Euc, ln(x+1)	17	0.927	0.525	0.687	0.829
Mod. Gower	23	0.895	0.532	0.708	0.841

< 500 μm

	m	prop.G	SS_{RES}	δ_1	correl
BC, sqrt	14	0.898	0.326	0.759	0.871
Jaccard	10	0.704	0.363	0.711	0.843
Euc, ln(x+1)	16	0.918	0.342	0.784	0.885
Mod. Gower	22	0.885	0.320	0.806	0.898

The best of these models were obtained on the basis of the Modified Gower measure, yielding canonical correlations of $r = 0.841$ and $r = 0.898$ for PC1.63 and PC1.500, respectively (Fig. 8). The canonical relationships obtained on the basis of the other dissimilarity measures, which were highly comparable when relative abundance information was included, are shown in Figure 9.

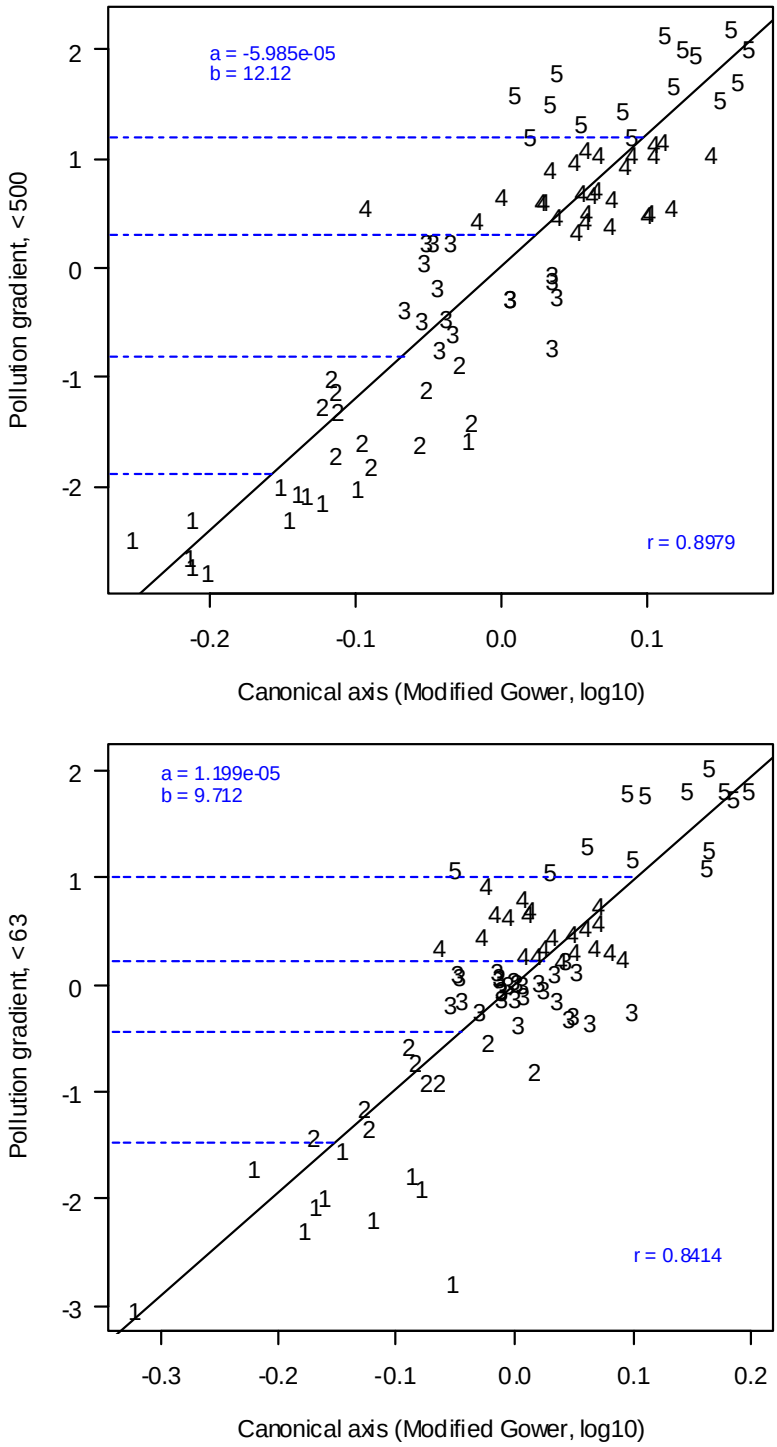


Fig. 8. CAP analyses of pollution gradients for the whole sample (< 500 µm, top) and the mud fraction (< 63 µm, bottom) versus faunal data (103 taxa) on the basis of the Modified Gower dissimilarity measure. Estimates of the intercept (*a*) and slope (*b*) of the predictive relationship as well as the correlation (*r*) are shown in blue. Also in blue are dotted lines demarcating the 5 pollution groupings.

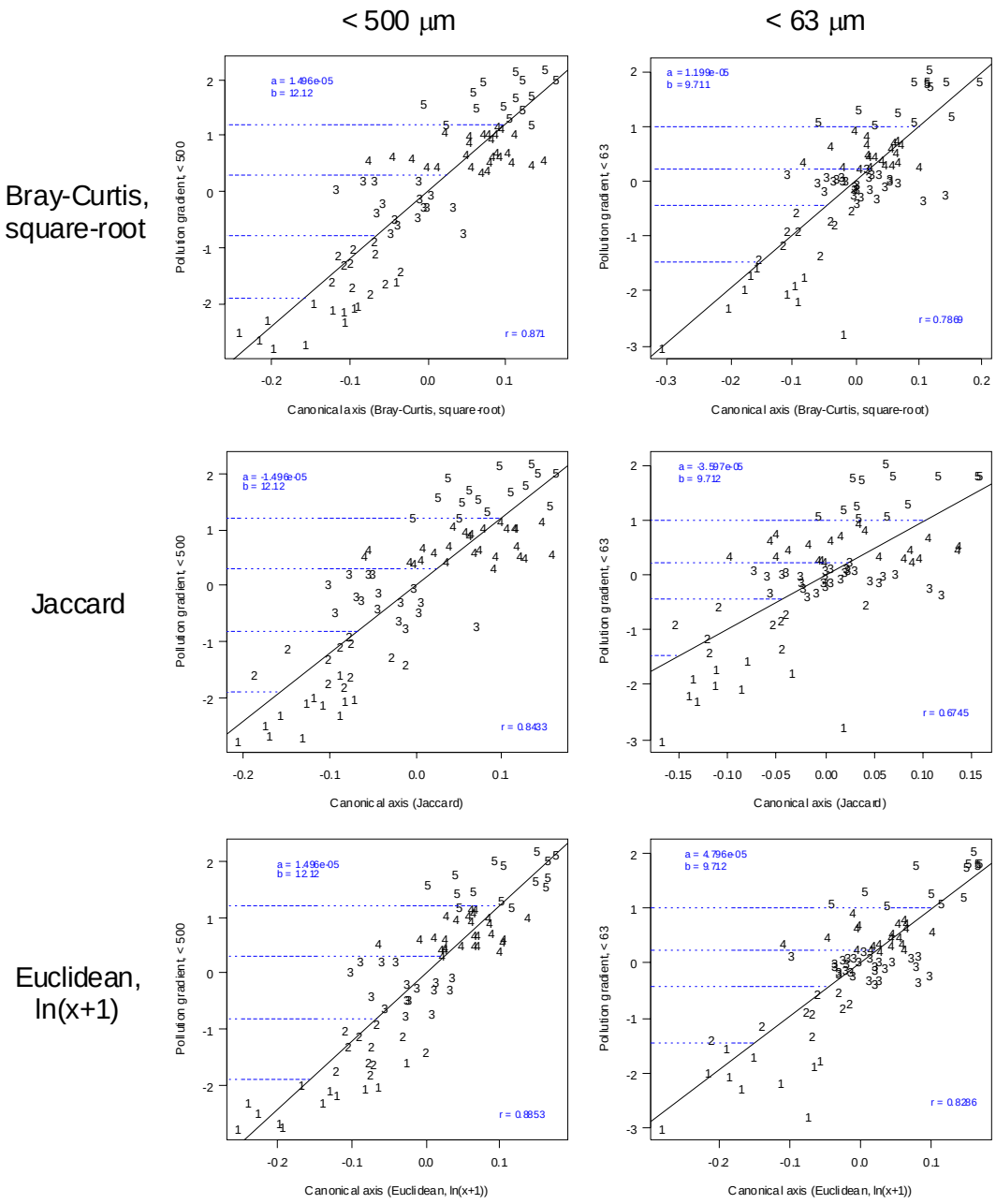


Fig. 9. CAP analyses of pollution gradients for the whole sample (< 500 μm, left) and the mud fraction (< 63 μm, right) versus faunal data (103 taxa) on the basis of each of three different dissimilarity measures. Estimates of the intercept (*a*) and slope (*b*) of the predictive relationship as well as the correlation (*r*) are shown in blue on each plot. Also in blue are dotted lines demarcating the 5 pollution groupings.

6.4 Predictive models using subsets of taxa

Scatterplots of individual taxa versus each of the pollution gradients (PC1.500 and PC1.63) are shown in Appendices 6 and 7, respectively. Only those taxa which occurred in at least 5% of the sites (i.e. 4 sites or more) were investigated in detail using these plots. The choice of a subset of taxa by eye using scatterplots is clearly suboptimal, because it is not possible to tell from these the degree to which individual taxa might overlap in their information content for individual sites. In addition, such choices also clearly depend on the eye of the beholder. Nevertheless, subsets of taxa were chosen in this manner (identified by asterisks in Appendices 6 and 7 and referred to hereafter as *scatter* subsets) for each of the pollution gradients: 14 taxa for PC1.500 and 7 taxa for PC1.63. The ecological subset (43 taxa) is given in Appendix 5. The scatter subset, the sensitivity subset (22 taxa) and the subsets obtained using BVSTEP or step-wise multiple regression with the BIC criterion (via DISTLM) are given for PC1.500 in Table 9 and for PC1.63 in Table 10. There was a reasonable amount of overlap in the particular subsets of taxa chosen to have the greatest relationship with each of the pollution gradient axes.

Our next step was to compare these subsets in terms of the strength of their relationship with the pollution gradients. This was done not by CAP but by calculating Spearman's rank correlation (rho = ρ) between dissimilarity matrices obtained using subsets of taxa (identified using BVSTEP and other methods) and the Euclidean distance matrix based on PC1.500. The values of ρ obtained using subsets were generally larger than that obtained using the whole set of 102 taxa (Table 11). The sensitivity, ecological and scatter subsets generally performed poorly. In contrast, the subsets obtained using BVSTEP achieved rank correlations well over 0.5 and the subset obtained using BVSTEP on Bray-Curtis for square-root transformed abundances achieved the best relationship with the PC1.500 pollution gradient, at $\rho = 0.589$ for a subset of 16 variables.

For the pollution gradient obtained using weak acid digestion on the mud fraction (PC1.63), the sensitivity, ecological and scatter subsets all did more poorly (had a weaker ρ value) than did the analysis using all taxa (Table 11). Subsets obtained using BVSTEP were better and the best relationship of those trialed was obtained using BVSTEP on the basis of Euclidean distance of ln(x+1)-transformed abundances (12 variables, $\rho = 0.536$).

We are interested here to actually model the pollution gradients directly, if possible, using the biotic data. For those subsets that were obtained on the basis of Euclidean distance for ln(x+1)-transformed abundances, each pollution gradient could be modeled directly using multiple regression. A univariate multiple regression analysis can be obtained using DISTLM when there is only one variable of interest (such as PC1.500) and Euclidean distance is used. In each case, the best model (in terms of the BIC criterion) was (not surprisingly) obtained using the subset which had been chosen using step-wise selection and the BIC criterion in the first place (in DISTLM), followed by the

BVSTEP subset. It was very interesting to note that the step-wise multiple regression subset on the basis of ln(x+1)-transformed abundances using BIC explained nearly 78% of the variance in PC1.500 with only 7 taxa and nearly 76% of the variance in PC1.63 with only 8 taxa (Table 12).

Table 9. Subsets of taxa versus PC1.500 identified considering the biology of the organisms (sensitivity), by using BVSTEP on the basis of various dissimilarity measures (as indicated), using visual examination of individual scatter plots (scatter), or by using DISTLM with the BIC criterion in a step-wise multiple regression of ln(x+1)-transformed variables. Names in blue were obtained in 3 or more of the 6 subsets.

Sensitivity	BVSTEP, BC, sqrt	BVSTEP, Euc, ln(x+1)	BVSTEP, MG	Scatter	DISTLM, BIC
Amphibola crenata	*Anthopleura aureoradiata*	*Anthopleura aureoradiata*	*Aglaophamus macroura*	*Amphibola crenata*	*Colurostylis* spp.
Amphipod other	*Cirolana sp.*	Bivalve unid.	*Anthopleura aureoradiata*	*Anthopleura aureoradiata*	*Helice, Hemi., Macrop.*
Aonides oxycephala	*Colurostylis* spp.	*Colurostylis* spp.	Chiton	*Austrovenus stutchburyi*	Nereidae
Aquilaspio aucklandica	*Edwardsia sp.*	*Crassostrea gigas*	*Colurostylis spp.*	*Colurostylis spp.*	Orbinidae
Aricidea sp.	*Haminoea zelandiae*	*Glycera* spp.	*Crassostrea gigas*	*Cominella glandiformis*	Phoxocephalidae
Arthritica bifurcata	Isopod other	*Helice, Hemi., Macrop.*	*Hiatula siliqua*	*Glycera* spp.	*Phyllodocid* spp.
Austrovenus stutchburyi	*Magelona sp.*	*Macomona liliana*	Lumbrineridae	*Helice, Hemi., Macrop.*	Sipunculid
Colurostylis spp.	*Minuspio sp.*	*Macroclymenella stewart.*	*Macroclymenella stewart.*	*Macomona liliana*	
Corophidae	Nereidae	*Mactra ovata*	*Magelona sp.*	*Mactra ovata*	
Cossura consimilis	Orbinidae	Nemertean	*Musculista senhousia*	Nereidae	
Exogoninae	*Paphies australis*	Nereidae	*Notomastus sp.*	*Notoacmea* spp.	
Glycera spp.	*Scolelepis* spp.	Orbinidae	Orbinidae	*Notomastus sp.*	
Helice, Hemi., Macrop.	Spionidae	*Paphies australis*	*Scolelepis* spp.	Phoxocephalidae	
Heteromastus filiformis	Sipunculid	*Scolelepis* spp.	Spionidae	*Scolecolepides benhami*	
Macomona liliana	*Trochodota dendyi*	Tanaidacea	Tanaidacea		
Macroclymenella stewart.	*Zeacumantus lutulentis*	*Waitangi brevirostris*	*Waitangi brevirostris*		
Mactra ovata		*Zeacumantus lutulentis*	*Zeacumantus lutulentis*		
Nemertean					
Nucula hartvigiana					
Paracalliope novizealandiae					
Phoxocephalidae					
Polydorid complex					

Table 10. Subsets of taxa versus PC1.63 identified considering the biology of the organisms (sensitivity), by using BVSTEP on the basis of various dissimilarity measures (as indicated), using visual examination of individual scatter plots (scatter), or by using DISTLM with the BIC criterion in a step-wise multiple regression of ln(x+1)-transformed variables. Names in blue were obtained in 3 or more of the 6 subsets.

Sensitivity	BVSTEP, BC, sqrt	BVSTEP, Euc, ln(x+1)	BVSTEP, MG	Scatter	DISTLM, BIC
Amphibola crenata	*Aglaophamus macroura*	*Aglaophamus macroura*	*Aglaophamus macroura*	*Amphibola crenata*	*Aglaophamus macroura*
Amphipod other	*Aricidea sp.*	*Crassostrea gigas*	*Magelona sp.*	*Glycera spp.*	*Cirolana sp.*
Aonides oxycephala	Barnacles	*Glycera spp.*	*Owenia fusiformis*	Lepidonotinae	Cirratulidae
Aquilaspio aucklandica	*Cirolana sp.*	Hesionidae	Spionidae	*Mactra ovata*	*Disconatus accolus*
Aricidea sp.	*Cominella glandiformis*	*Mactra ovata*	*Waitangi brevirostris*	Nereidae	*Glycera spp.*
Arthritica bifurcata	*Felaniella zelandica*	*Magelona sp.*		Phoxocephalidae	Lepidonotinae
Austrovenus stutchburyi	*Glycera spp.*	*Musculista senhousia*		*Scolecolepides benhami*	Nereidae
Colurostylis spp.	*Halicarcinus* spp.	Mysidacea			Phoxocephalidae
Corophidae	*Magelona sp.*	*Scolecolepides benhami*			
Cossura consimilis	Mantis shrimp	Spionidae			
Exogoninae	Mysidacea	Sipunculid			
Glycera spp.	Nereidae	*Waitangi brevirostris*			
Helice, Hemi., Macrop.	Orbinidae				
Heteromastus filiformis	*Paphies australis*				
Macomona liliana	Phoxocephalidae				
Macroclymenella stewart.	*Phyllodocid* spp.				
Mactra ovata	Polynoid				
Nemertean	Spionidae				
Nucula hartvigiana	Syllinae				
Paracalliope novizealandiae	*Waitangi brevirostris*				
Phoxocephalidae					
Polydorid complex					

Table 11. Relationship between the dissimilarity matrix obtained using each of several subsets of taxa versus the Euclidean distance matrix obtained from either PC1.500 or PC1.63, as indicated. The set labeled "All" indicates analysis of the full set of 102 variables. ρ = the Spearman rank correlation coefficient, no. vars = the number of variables in the subset.

	PC1.500		PC1.63	
Subset	no. vars	ρ	no. vars	ρ
All, BC, sqrt	102	0.204	102	0.374
All, Euc, ln(x+1)	102	0.123	102	0.297
All, MG	102	0.202	102	0.347
Sensitivity, BC, sqrt	22	0.284	22	0.163
Sensitivity, Euc, ln(x+1)	22	0.230	22	0.055
Sensitivity, MG	22	0.299	22	0.171
Ecological, BC, sqrt	43	0.343	43	0.204
Ecological, Euc, ln(x+1)	43	0.282	43	0.121
Ecological, MG	43	0.327	43	0.194
BVSTEP, BC, sqrt	16	0.589	20	0.438
BVSTEP, Euc, ln(x+1)	17	0.572	12	0.536
BVSTEP, MG	17	0.568	5	0.454
Scatter, BC, sqrt	14	0.479	7	0.307
Scatter, Euc, ln(x+1)	14	0.446	7	0.287
Scatter, MG	14	0.446	7	0.296
DISTLM, BIC, ln(x+1)	7	0.411	8	0.260

Table 12. Results of DISTLM models (multiple regression) to explain variation in either PC1.500 or PC1.63 on the basis of each of several sets of taxon variables. No. vars = the number of variables included in the regression, % explained = percentage of variation in the response variable (in each case) explained by each set of regression variables, SS_{Res} = the residual sum of squares, BIC = information criterion which takes into account the number of variables in the model (smaller values indicate a better model).

PC1.500

Subset	no. vars	% explained	SS_{Res}	BIC
Sensitivity, Euc, ln(x+1)	22	70.26%	43.680	46.66
Ecological, Euc, ln(x+1)	43	82.47%	25.751	96.13
BVSTEP, Euc, ln(x+1)	17	78.17%	32.067	-0.35
Scatter, Euc, ln(x+1)	14	75.16%	36.490	-3.07
DISTLM, BIC, ln(x+1)	7	77.66%	32.817	-42.42

PC1.63

Subset	no. vars	% explained	SS_{Res}	BIC
Sensitivity, Euc, ln(x+1)	22	60.32%	37.433	34.15
Ecological, Euc, ln(x+1)	43	80.33%	18.552	69.58
BVSTEP, Euc, ln(x+1)	12	71.41%	26.967	-36.35
Scatter, Euc, ln(x+1)	7	55.18%	42.277	-21.91
DISTLM, BIC, ln(x+1)	8	75.88%	22.750	-67.71

Note that there is a difference between calculating a correlation between ranks of distance matrices (Table 11) and direct regression modeling of the variance in pollution gradients (Table12). The regression models were only done for subsets that had been

obtained originally within a Euclidean distance framework. For each of PC1.500 and PC1.63, three approaches yielded the best subsets of variables for modelling, according to either the DISTLM analysis or Spearman's rho (ρ). These were: (i) BVSTEP on Bray-Curtis dissimilarities of square-root transformed abundances, (ii) BVSTEP on Euclidean distances of ln(x+1)-transformed abundances or (iii) step-wise selection using BIC on ln(x+1)-transformed abundances. Therefore, CAP analyses were done to explore these subsets further.

A summary of these CAP analyses is shown in Table 13 and accompanying plots are shown in Fig. 10. For PC1.63, the strongest canonical correlation was obtained using the subset identified using BIC with step-wise regression on ln(x+1)-transformed variables. For PC1.500, the subset chosen using the BVSTEP algorithm on Bray-Curtis dissimilarities of square-root transformed abundances and the subset obtained using DISTLM step-wise regression on ln(x+1)-transformed variables performed well. The predictive capacity of these subsets also appears to be the best, as they achieved the smallest leave-one out residual sum of squares. It is rather impressive that the step-wise regression subsets obtained using DISTLM performed so well, especially given that these subsets had only 7 or 8 variables (for PC1.500 or PC1.63, respectively). Also interesting is the fact that the CAP models obtained using subsets were better than those obtained using the full set of data in the case of PC1.63. This was not the case, however, for PC1.500, for which prediction using the whole complement of species performed better than CAP models using only subsets (compare results in Table 13 with those in Table 8).

Table 13. Summary of CAP analyses to model pollution gradients (obtained from PCA on log metal concentrations) using each of 3 different subsets of taxa from the faunal data, as indicated. Results are given separately for the metal concentrations from the mud fraction (< 63 μm) and from the total sample (< 500 μm). *p* = the number of variables in the subset, *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, SS_{RES} = the leave-one-out residual sum of squares, δ_1 is the squared canonical correlation for the canonical axis, correl = the correlation between the canonical axis and the pollution gradient.

< 63 μm

	p	*m*	prop.G	SS_{RES}	δ_1	correl
BVSTEP, BC, sqrt	20	10	0.996	0.513	0.672	0.820
BVSTEP, Euc, ln(x+1)	12	12	1.000	0.527	0.714	0.845
DISTLM, BIC, ln(x+1)	8	8	1.000	0.320	0.759	0.871

< 500 μm

	p	*m*	prop.G	SS_{RES}	δ_1	correl
BVSTEP, BC, Sqrt	16	3	0.691	0.272	0.748	0.865
BVSTEP, Euc, ln(x+1)	17	6	0.865	0.306	0.740	0.860
DISTLM, BIC, ln(x+1)	7	4	0.948	0.296	0.733	0.856

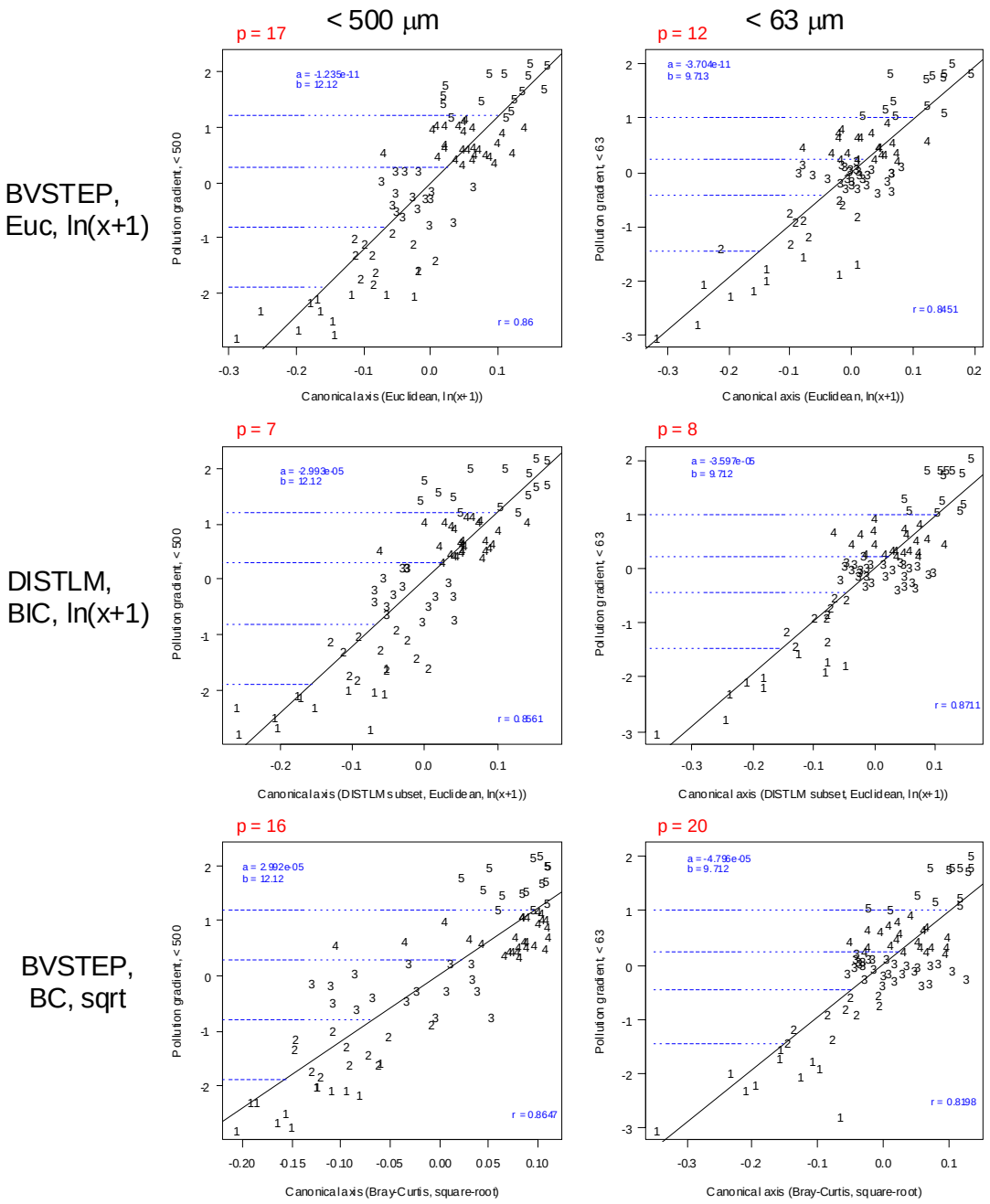


Fig. 10. CAP analyses of pollution gradients for the whole sample (< 500 µm, left) and the mud fraction (< 63 µm, right) versus fauna for each of three different subsets, as indicated. Estimates of the intercept (*a*) and slope (*b*) of the predictive relationship as well as the correlation (*r*) are shown in blue on each plot. Also in blue are dotted lines demarcating the 5 pollution groupings. The number of variables (i.e., the number of taxa, *p*) in each subset are indicated in red.

6.5 Refinements using physical variables

For many of the physical variables, a log transformation worked virtually as well as or better than the optimal power transformation to render the data approximately normal (Appendix 4). There were two exceptions to this: for the coarse sand fraction, the best transformation was obtained using a power of –0.79 and for the silt and clay fraction, the best transformation was obtained using a power of 0.37 (Appendix 4). Thus, in what follows, all of the physical variables were transformed to logs except for these two variables, which were transformed using their corresponding optimal power transformation, as indicated.

A PCA of the transformed and normalised physical variables showed that most of the variation among sites in terms of these physical characteristics was the contrast between the silt and clay fraction on the one hand and the coarse sand fraction on the other (Fig. 11). The degree of exposure (CWE and FWE) also contributed a reasonable amount towards the total variation in physical characteristics among sites (Fig. 11, Table 14).

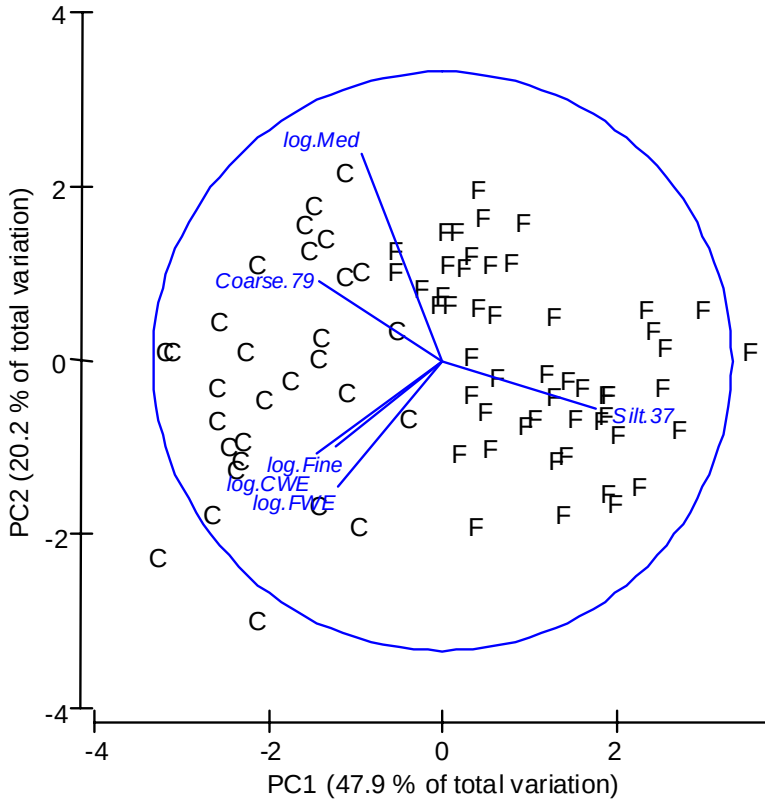


Fig. 11. PCA of 6 optimally transformed and normalised physical variables at each of 81 sites. Letters indicate the 2 groups (F = finer sediments, less exposed and C = coarser sediments, more exposed) obtained using *k*-means partitioning on the basis of Euclidean distances.

Table 14. Coefficients (also called "weights" or "eigenvectors") and the percentage of the total variance explained by each of the principal component axes from a PCA on normalised, optimally-transformed physical variables. The mean and standard deviation (SD) used for the normalisation of each variable is shown. Most variables were transformed using natural logs (or ln(x+1) if there were zeros), but coarse sand was transformed using a power transformation of -0.79 and silt and clay was transformed using a power transformation of 0.37.

			PC1	PC2	PC3	PC4	PC5	PC6
			47.9%	20.2%	13.3%	8.5%	7.2%	2.9%
Variable	Mean	SD	Eigenvectors					
Coarse sand'	0.526	0.365	-0.426	0.275	-0.172	0.843	0.035	-0.028
ln(Medium sand + 1)	2.257	1.444	-0.283	0.719	0.229	-0.342	-0.093	-0.474
ln(Fine sand)	3.162	0.871	-0.368	-0.298	-0.739	-0.246	-0.154	-0.382
Silt and Clay'	7.251	3.588	0.531	-0.164	0.080	0.310	0.069	-0.764
ln(FWE)	0.775	1.074	-0.365	-0.433	0.523	0.085	-0.618	-0.127
ln(CWE + 1)	1.963	1.149	-0.434	-0.322	0.304	-0.089	0.761	-0.168

Hierarchical agglomerative group-average cluster analysis of the Euclidean distance matrix based on the normalised transformed physical variables did not show very clear groupings of sites (Fig. 12). An arbitrary slice at a Euclidean distance value of 3.9 (which was also made to avoid obtaining a single site as a separate group on its own) yielded 2 groups. SIMPROF analysis indicated that there were 5 distinct groups (Fig. 12). Analysis by *k*-means partitioning suggested that the best grouping structure was achieved for $k = 2$ groups, with the highest value of $CH_k = 293.85$ (Table 15).

The SIMPROF 5-group solution seemed slightly better than the others according to the ANOSIM test statistic, whereas the *k*-means 2-group solution seemed best according to the PERMANOVA test statistic, which is based directly on inter-point distances (Table 16). Due to the fact that the SIMPROF 5-group solution had large discrepancies in the number of sites per group (one group had only one sample and another group had only three samples, see Fig. 12) and the *k*-means solution yielded groups that were clearly distinct in the PCA (Fig. 11), the 2-group solution was retained for further analysis.

Table 15. Values of the Calinski-Harabasz criterion (CH_k) for the K-means partitioning solutions from K = 2-12 groups for the normalised transformed physical variables.

No. groups (K)	CH_k
2	293.85
3	202.06
4	159.60
5	132.82
6	114.47
7	100.97
8	97.10
9	88.44
10	85.27
11	74.25
12	75.23

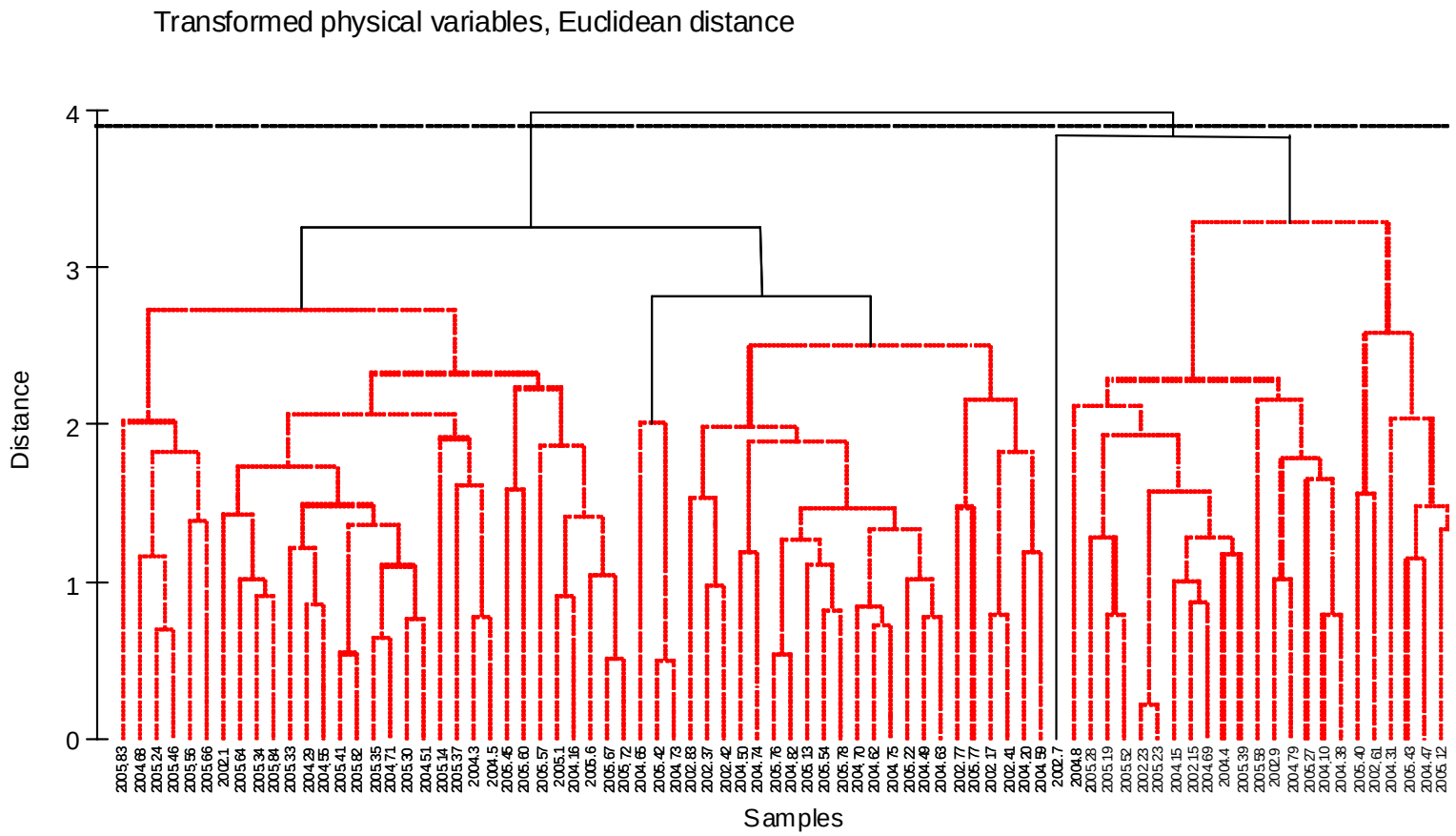


Fig. 12. Hierarchical agglomerative group-average cluster analysis among samples on the basis of optimally transformed physical variables using Euclidean distances. The splits in red were not statistically significant, by SIMPROF, which indicated 12 groups of samples. The single horizontal line indicates an arbitrary slice at a distance value of 3.9 which yielded 2 groups. Labels indicate the year and site number (as per Appendix 1) for each sampling unit.

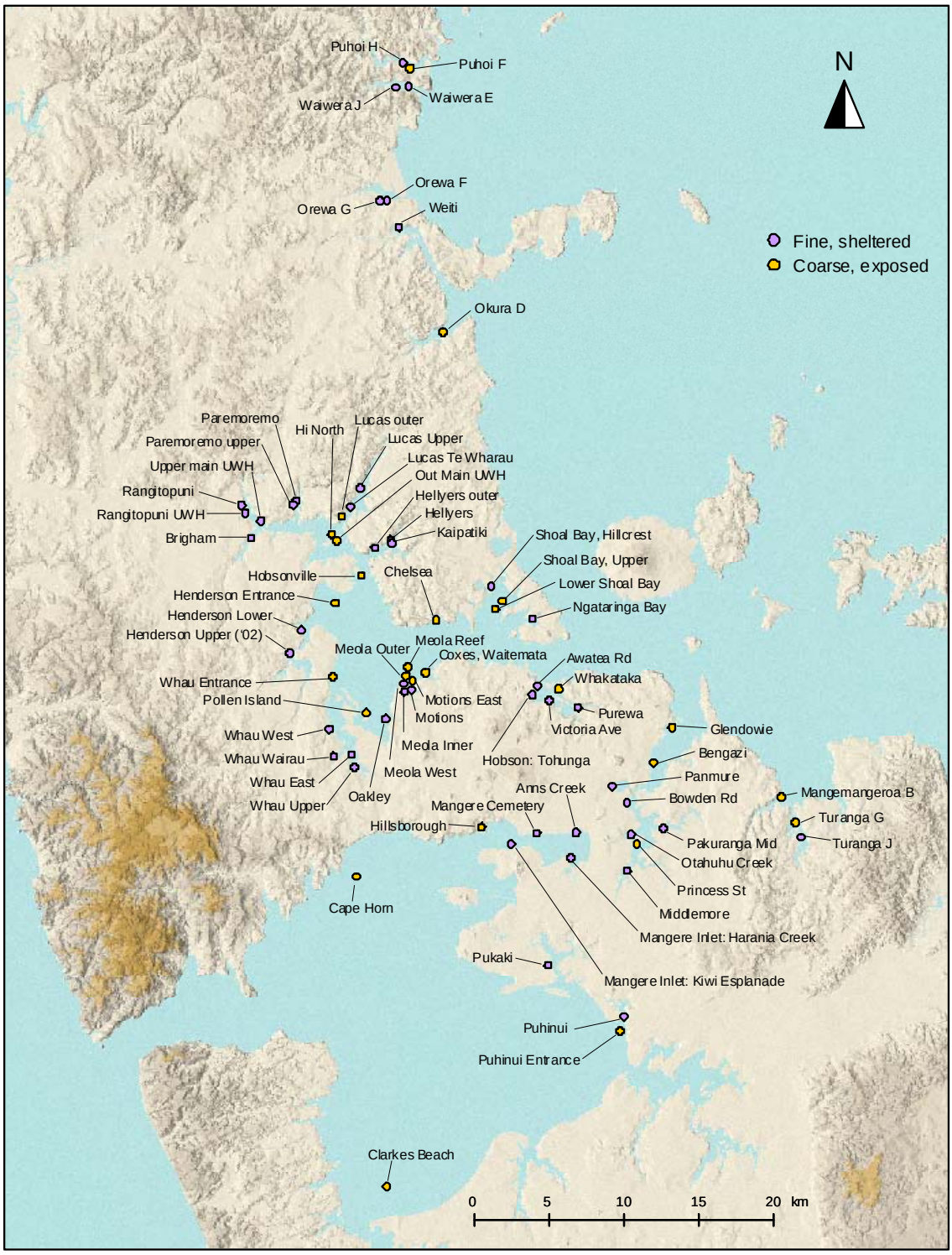


Fig. 13. Map of the Auckland region showing each of the 81 sites in terms of two groups on the basis of *k*-means partitioning of the six physical variables (F = relatively fine sediments and lesser exposure, C = relatively coarse sediments with greater exposure).

Table 16. Values of the ANOSIM *R*-statistic and the PERMANOVA *F*-statistic for the transformed normalized physical variables with respect to each of the grouping solutions.

	k	ANOSIM *R*	PERMANOVA *F*
k-means	2	0.669	43.071
Slice	2	0.640	37.957
SIMPROF	5	0.741	22.953

The two groups identified by *k*-means are interpretable as: a group consisting of samples from sites having relatively finer sediments and lesser wind exposure ("group F"), and a second group consisting of samples from sites having relatively coarser sediments and greater wind exposure ("group C"), as evidenced by their positions in the PCA plot (Fig. 11). There were 51 samples in group F and 30 samples in group C. A map identifying each site by reference to this two-group *k*-means solution on the physical variables is provided in Fig. 13.

Table 17. Summary of CAP analyses relating biotic assemblages to pollution gradients based on either the whole sample (< 500 μm) or on the mud fraction (< 63 μm) done separately on separate datasets obtained by splitting the data according to k-means partitioning of physical variables into two groups. Set F = fine sediments, more sheltered (51 samples). Set C = coarser sediments, more exposed (30 samples). Table headings are as given for Table 13 above.

<500 μm

	Set	*m*	prop.G	SS_{RES}	δ_1	correl
BC, sqrt	F	9	0.845	0.370	0.726	0.852
Euc, ln(x+1)	F	6	0.759	0.407	0.669	0.818
Mod. Gower	F	9	0.733	0.414	0.699	0.836
BC, sqrt	C	8	0.822	0.545	0.619	0.787
Euc, ln(x+1)	C	6	0.747	0.643	0.572	0.756
Mod. Gower	C	7	0.669	0.513	0.643	0.802

< 63 μm

	Set	*m*	prop.G	SS_{RES}	δ_1	correl
BC, sqrt	F	5	0.668	0.585	0.497	0.705
Euc, ln(x+1)	F	4	0.641	0.660	0.478	0.692
Mod. Gower	F	6	0.613	0.622	0.489	0.699
BC, sqrt	C	17	0.991	0.162	0.970	0.985
Euc, ln(x+1)	C	15	0.948	0.193	0.943	0.971
Mod. Gower	C	14	0.886	0.260	0.909	0.953

Next, CAP analyses were done separately for each group on the basis of each of three different dissimilarity measures relating the biota to each of the identified pollution

gradients: PC1.500 or PC1.63 (Table 17). The best relationship for the more sheltered habitats (F) was with the metal concentrations measured from the whole sample (<500 microns), whereas the best relationship for the more exposed habitats (C) was with the metal concentrations measured from the weak acid digestion of the mud fraction (<63 microns) Fig. 14). Although similar results were obtained using any of the three dissimilarity measures, those done using Bray-Curtis on square-root transformed abundances tended to achieve the highest canonical correlations (Fig. 14). Indeed, when the analysis was based on Bray-Curtis from square-root transformed data, the squared canonical correlation between the biota from samples having coarser sediments (C) and PC1.63 was 0.970, the strongest relationship obtained for any of the canonical models examined thus far. Such high correlations can be a little deceptive, however, as these values also depend on the number of PCO axes (*m*) that were used in the model, which was fairly large (*m* 14) for those analyses of biota from samples in group C versus PC1.63. Nevertheless, the value of *m* in all cases was chosen so as to minimize the leave-one-out residual sum of squares, which in this case is substantially smaller for these models than for the models of group C samples versus PC1.500.

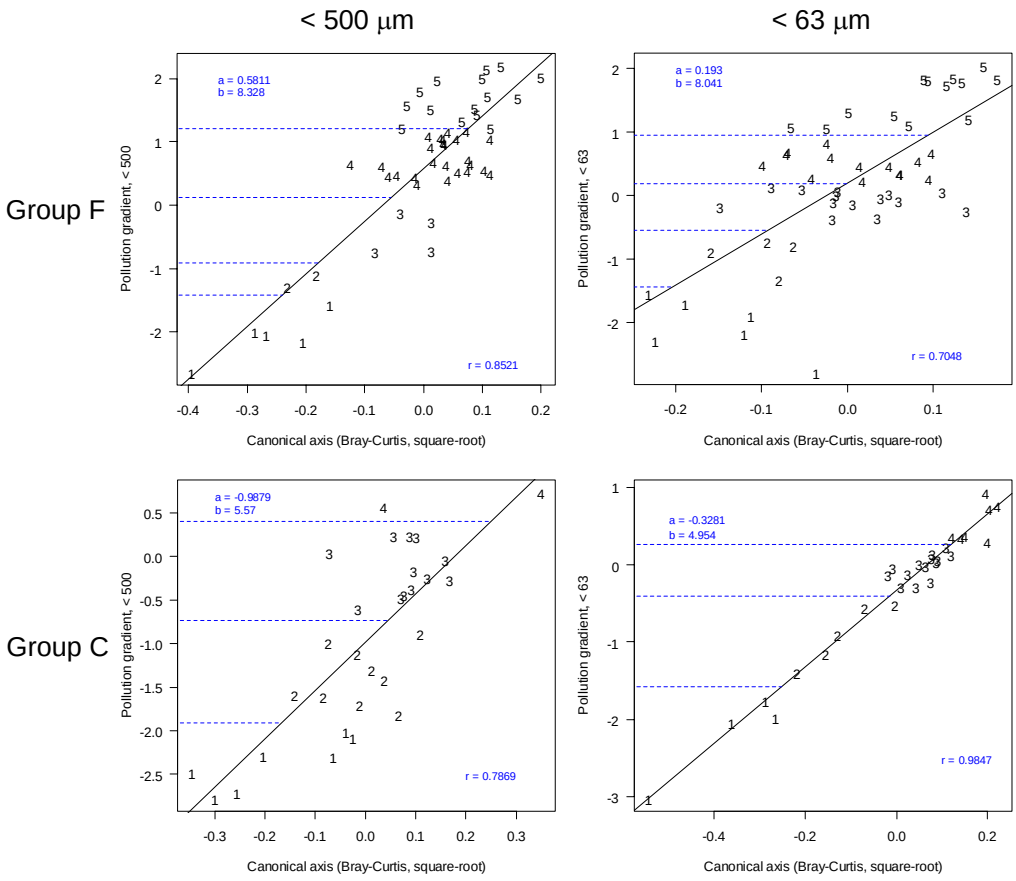


Fig. 14. CAP analyses of pollution gradients for the whole sample (< 500 µm, left) and the mud fraction (< 63 µm, right) versus fauna for each of 2 groups (F = finer sediments, less exposed and C = coarser sediments, more exposed) obtained using *k*-means partitioning on the basis of Euclidean distances, as indicated. Estimates of the intercept (*a*) and slope (*b*) of the predictive relationship as well as the correlation (*r*) are shown in blue on each plot. Also in blue are dotted lines demarcating the 5 pollution groupings.

Importantly, the partitioning of the data on the basis of the physical variables resulted in better CAP models of PC1.63 for samples in group C. This suggests that biota occurring in such habitats (i.e. with coarser sediments and greater exposure) are more sensitive to pollutants measured in the < 63 μm fraction. The relationship of the biota to PC1.500 was not, however, necessarily improved by this partitioning for samples in group F, but was largely unchanged. Prediction of ecosystem health for these sites versus PC1.500 would be expected to be marginally better using the entire dataset than using only sites from group F. Independent validation is required, however, to determine the best overall modeling approach across the region. These analyses suggest that one possible approach may be to model sites in group C using PC1.63 and sites in group F using PC1.500.

For comparison, we also considered the previously defined habitat definitions of Settling Zone and Outer Zone (ARC 2002), which correspond roughly to group F and group C, respectively (Appendix 8). CAP analyses of PC1.500 and PC1.63 were done separately for each of the SZ and OZ groups, and results were similar to those obtained for groups F and C (compare Table 17 with Appendix 9). However, CAP models of PC1.63 versus group C were stronger than those of PC1.63 versus OZ.

Subsets of variables that might be driving relationships between the biota and the pollution gradients (either PC1.500 for group F or PC1.63 for group C) were identified using (i) BVSTEP on BC dissimilarities of sqrt-transformed data; (ii) BVSTEP on Euclidean distances of ln(x+1)-transformed data and (iii) step-wise selection of ln(x+1)-transformed data using DISTLM. Results of CAP and RELATE analyses for each of these subsets are given in Table 18 and each subset is identified in Table 19.

Table 18. Summary of CAP analyses for subsets obtained using separate groups of samples identified from physical variables: no. vars = the number of variables in the subset, ρ = Spearman's rank correlation with the Euclidean distance matrix from the pollution gradient. Other headings are as given in Table 13 above.

Group C, < 63 μm

Subset	*no. vars*	ρ	*m*	prop.G	SS_{RES}	δ_1	correl
BVSTEP, BC, sqrt	22	0.690	5	0.868	0.245	0.836	0.914
BVSTEP, Euc, ln(x+1)	9	0.811	3	0.990	0.538	0.654	0.809
DISTLM, BIC, ln(x+1)	28	0.340	6	0.862	0.212	0.853	0.924

Group F, < 500 μm

Subset	*no. vars*	ρ	*m*	prop.G	SS_{RES}	δ_1	correl
BVSTEP, BC, sqrt	16	0.569	7	0.989	0.298	0.812	0.901
BVSTEP, Euc, ln(x+1)	10	0.648	2	0.780	0.516	0.507	0.712
DISTLM, BIC, ln(x+1)	6	0.275	5	0.996	0.331	0.749	0.865

For group C, none of the subsets of variables achieved canonical correlations with PC1.63 as high as those obtained using the whole set of variables (compare the values

in Table 17 and Table 18). However, for group F, both the BVSTEP subset based on Bray-Curtis (16 variables, correlation = 0.901) and the subset based on DISTLM of ln(x+1)-transformed abundances (6 variables, correlation = 0.865) were improvements on the model of the group F sites that included all variables. These were also comparable to the models for PC1.500 that included all of the sites (cf. Table 8).

Table 19. Subsets of taxa for group F samples versus PC1.500 and for group C samples versus PC1.63 identified by using BVSTEP on the basis of various dissimilarity measures (as indicated), or by using DISTLM with the BIC criterion in a step-wise multiple regression of ln(x+1)-transformed variables.

GROUP F *versus* PC1.500			GROUP C *versus* PC1.63		
BVSTEP, BC, sqrt	BVSTEP, Euc, ln(x+1)	DISTLM, BIC	BVSTEP, BC, sqrt	BVSTEP, Euc, ln(x+1)	DISTLM, BIC
Aonides oxycephala	*Aonides oxycephala*	*Aonides oxycephala*	*Aglaophamus macroura*	*Aglaophamus macroura*	*Anthopleura aureoradiata*
Nereidae	*Scolelepis* spp.	*Austrovenus stutchburyi*	*Alpheus sp.*	*Alpheus sp.*	*Aonides oxycephala*
Scolelepis spp.	Bivalve unid.	*Glycera* spp.	*Amphibola crenata*	Anthuridae	*Aquilaspio aucklandica*
Anthopleura aureoradiata	*Colurostylis* spp.	*Heteromastus filiformis*	Anthuridae	*Crassostrea gigas*	*Aricidea sp.*
Bivalve unid.	*Magelona sp.*	Nereidae	*Aricidea sp.*	Hesionidae	*Asychis amphiglypta*
Colurostylis spp.	*Musculista senhousia*	*Scolelepis* spp.	*Cirolana sp.*	*Magelona sp.*	Barnacles
Edwardsia sp.	Orbinidae		*Cominella adspersa*	*Minuspio sp.*	*Cominella glandiformis*
Hesionidae	Spionidae		*Cominella glandiformis*	Platyhelminth	*Cossura consimilis*
Isopod other	*Zeacumantus lutulentis*		Corophidae	Sipunculid	*Crassostrea gigas*
Magelona sp.	*Haminoea zelandiae*		*Crassostrea gigas*		*Cyclaspis thomsoni*
Mantis shrimp			*Felaniella zelandica*		*Helice, Hemi., Macrop.*
Musculista senhousia			Hesionidae		*Hiatula siliqua*
Orbinidae			*Hiatula siliqua*		Isopod other
Spionidae			*Magelona sp.*		*Macroclymenella stewart.*
Zeacumantus lutulentis			*Minuspio sp.*		*Mactra ovata*
Zediloma subrostrata			Mysidacea		*Magelona sp.*
			Nereidae		Nereidae
			Phyllodocid spp.		*Notoacmea spp.*
			Polynoid		*Nucula hartvigiana*
			Sipunculid		Opistobranch
			Syllinae		Orbinidae
			Xymene sp.		*Paphies australis*
					Platyhelminth
					Polydorid complex
					Pontophilus australis
					Sipunculid
					Trochodota dendyi
					Xymene sp.

6.6 Validation

The position of each of the validation sites on each of the pollution gradients (i.e. for the total sample, PC1.500, and for the < 63 μm fraction, PC1.63) based on their actual measured metal concentrations are shown in Table 20. There was quite a spread of samples in the validation set in terms of levels of contaminants, with several sites in each of the 5 pollution groupings obtained for both gradients. It is interesting to note also that some sites differed rather strongly in their values between the two gradients. For example, both Newmarket (site 44) and Coxs (site 11) were quite low (group 2) along pollution axis PC1.500, but were quite high (group 5) along pollution axis PC1.63. A visual representation of the pollution levels in terms of the total sample (PC1.500) and in terms of the mud fraction (PC1.63) for all of the sites in this study are shown on the maps in Fig. 15 and Fig. 16, respectively.

Table 20. Allocation to a physical grouping and positions for each validation site along each pollution gradient, along with its allocation to a rank state of benthic health in each case.

Year	Site no.	Site name	Physical group	PC1.500	rank.500	PC1.63	rank.63
2002	2	Auckland Airport	C	-2.085	1	-1.922	1
2005	11	Coxs	C	-1.103	2	1.406	5
2005	17	Henderson Upper	F	1.329	5	0.770	4
2005	18	Herald Island	F	-0.438	3	-0.180	3
2005	21	Hobson - Purewa Bridge	F	0.776	4	0.172	3
2004	25	Kendalls	C	-1.744	1	-0.730	2
2006	26	Little Shoal Bay	C	-1.134	2	-0.099	3
2004	32	Mangemangeroa E	C	-0.700	3	-0.945	2
2005	36	Mangere Inlet: Tararata Creek	F	0.816	4	-0.188	3
2005	44	Newmarket	C	-0.991	2	1.033	5
2004	48	Okura J	C	-2.095	1	-2.198	1
2005	53	Pakuranga	F	1.600	5	1.534	5
2002	80	Whau Entrance, WHO A	C	-1.583	2	0.170	3
2005	81	Whau Lower	F	1.255	5	0.756	4

Section 4.5 above describes models which include physical information. Specifically, sites belonging to group C (relatively coarse sediments, greater exposure) were modeled with PC1.63 and sites belonging to group F (relatively fine sediments, less exposure) were modeled separately with PC1.500. To perform this kind of "split" modeling with the validation sites, it was first necessary to allocate each validation site to one of the two physical groups (C or F). This was done by placing each site into the two-dimensional PCA diagram shown in Fig. 11 using the eigenvector coefficients, means and standard deviations shown in Table 14 above. The result is shown in Fig. 17 below. From this, each validation site was easily allocated into one of the two groups,

as shown in Table 20. Although site 11 was placed into group C, it was perhaps marginal. The allocations shown in Table 20 above and on the two-dimensional PCA plot were also obtained exactly, however, when a CAP discriminant analysis on the physical data was used instead (results not shown here). There were 8 validation sites allocated to group C and 6 allocated to group F.

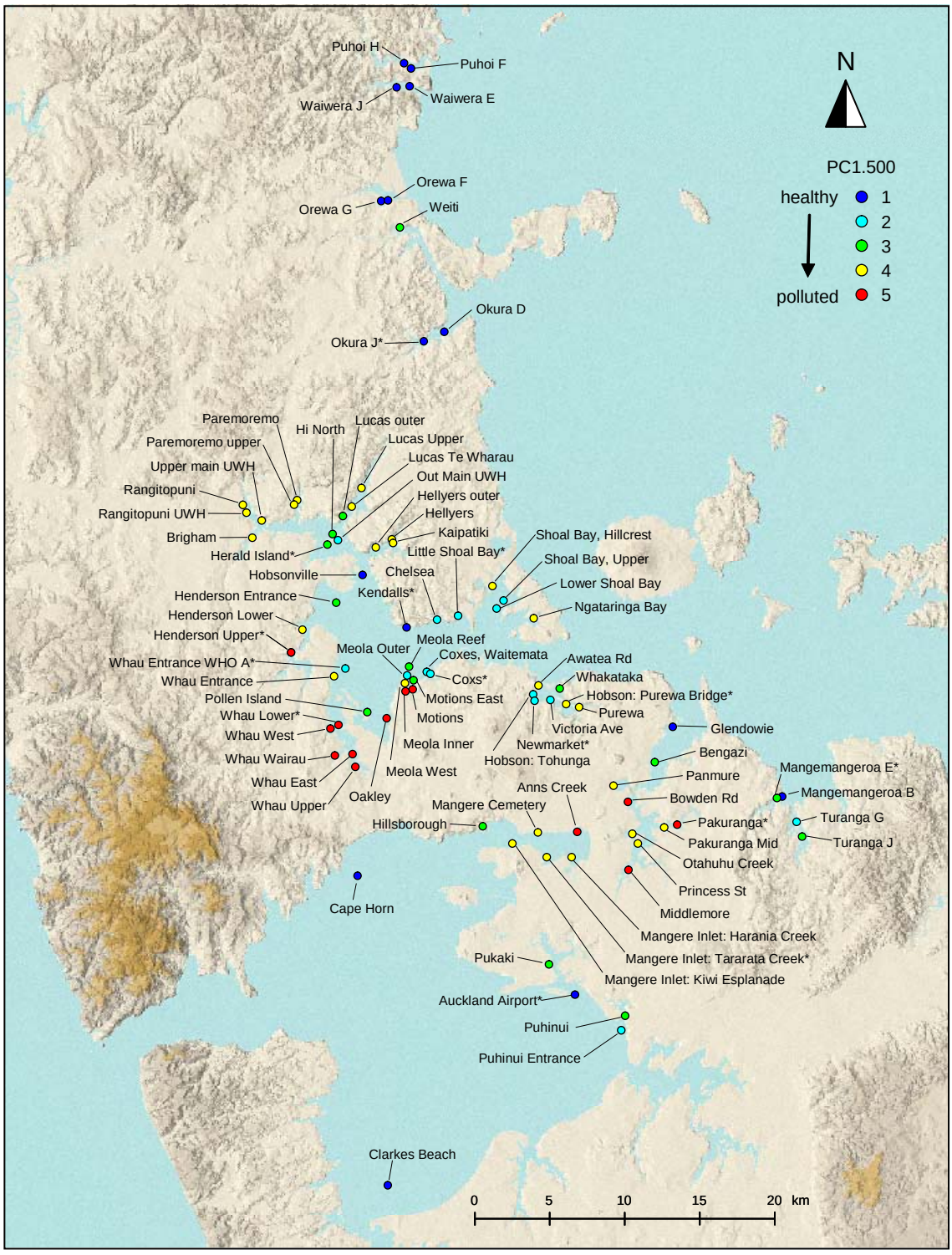


Fig. 15. Map of the Auckland region showing pollution groupings on the basis of *k*-means partitioning (1 = healthy, 5 = polluted) of the first principal component of log metal concentrations (Cu, Pb and Zn) for the whole sample (< 500 μm) at each of 95 sites, including validation sites (denoted by *).

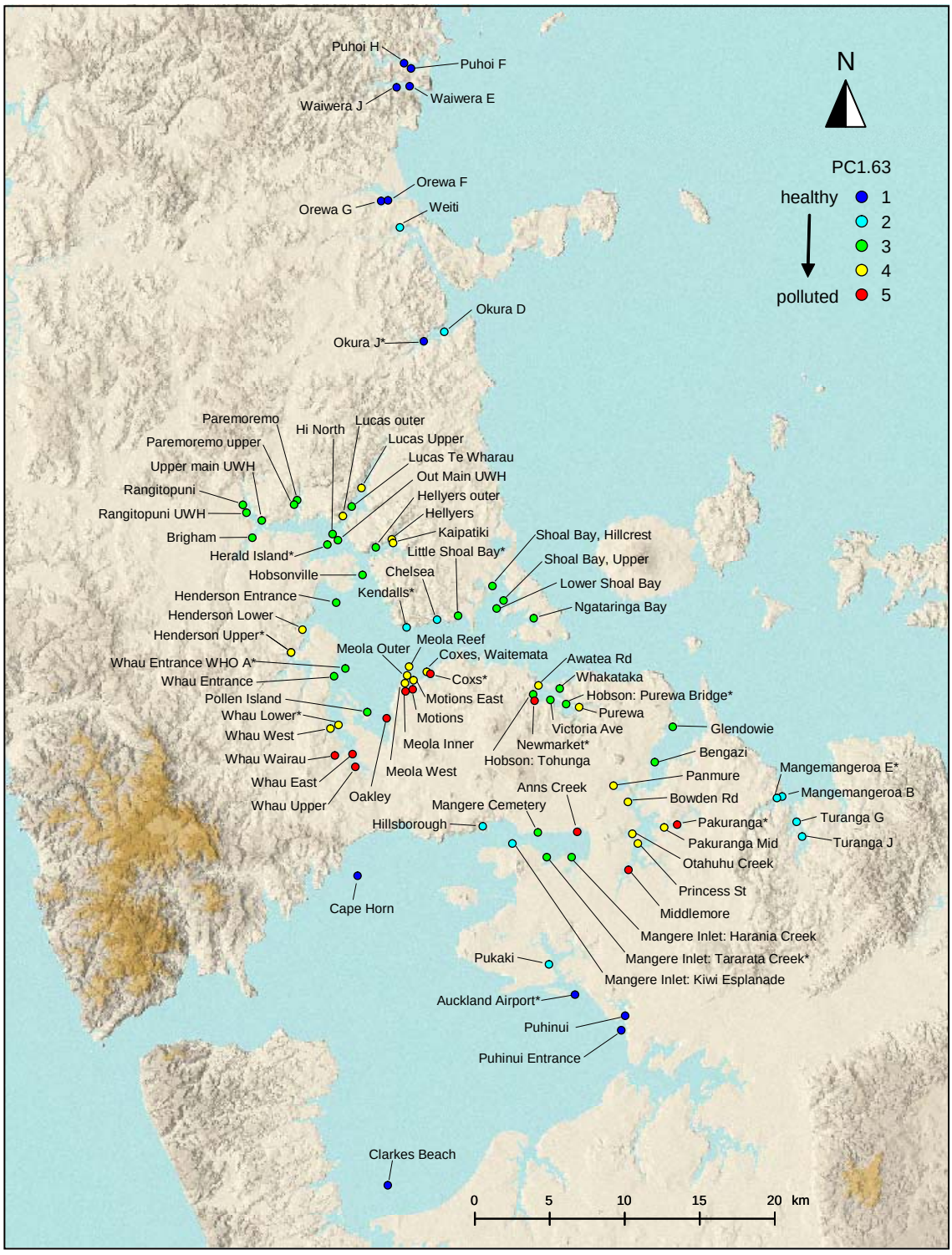


Fig. 16. Map of the Auckland region showing pollution groupings on the basis of *k*-means partitioning (1 = healthy, 5 = polluted) of the first principal component of log metal concentrations (Cu, Pb and Zn) for the mud fraction (< 63 μm) at each of 95 sites, including validation sites (denoted by *).

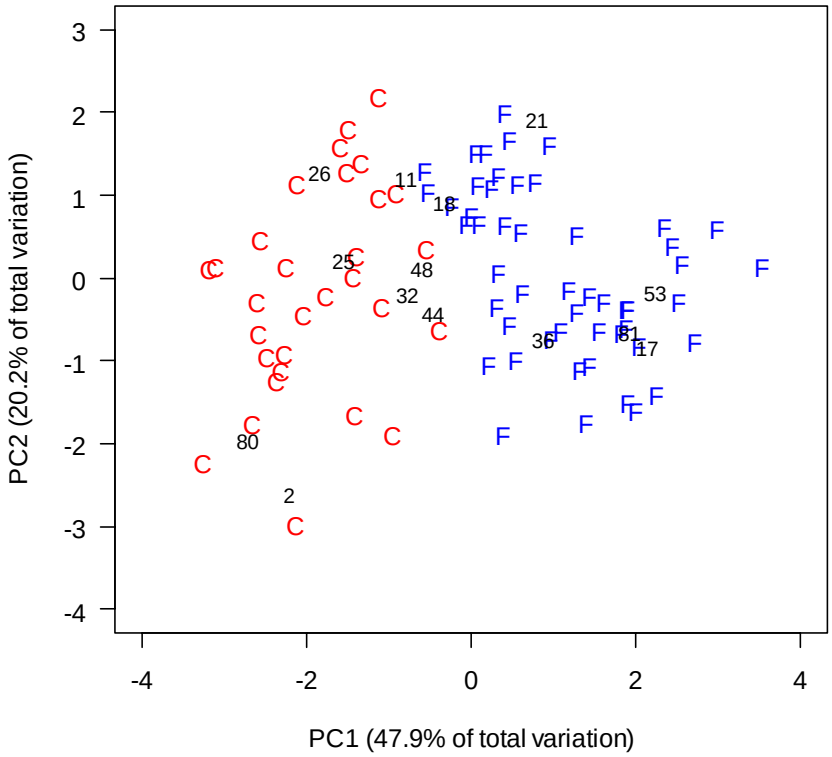


Fig. 17. PCA of 6 optimally transformed and normalised physical variables at each of 81 sites. Letters indicate the 2 groups (F = finer sediments, less exposed and C = coarser sediments, more exposed) obtained using *k*-means partitioning on the basis of Euclidean distances. Numbers corresponding to each validation site (listed in Table 20) are superimposed in black.

Not all of the models were considered as part of the validation process: only those which showed some clear promise based on the modeling exercise itself or were of interest for a priori reasons (i.e. the sensitivity and ecological subsets) were pursued further. This did result, however, in quite a large number of potential models. The residual sum of squares calculated as the sum of squared deviations of predicted values from the actual values along each pollution gradient (SS_{RES}) are shown for all relevant models in Table 21.

Interestingly, the best models for PC1.500 were obtained from those that used only subsets of taxa. Overall, the best two models were the 16 taxa of the BVSTEP subset obtained on the basis of Bray-Curtis dissimilarities of square-root transformed abundances and the *a priori* chosen sensitivity subset of 22 taxa (Table 21). Even though the sensitivity subset did not show a particularly strong canonical correlation compared to some of the other subsets, it performed well for prediction. The subset obtained using the BIC criterion from step-wise selection of ln(x+1)-transformed abundances with DISTLM also performed surprisingly well, especially given that it included only 7 taxa.

Table 21. Models of benthic ecosystem health and their associated sum of squared deviations of predicted from actual values along relevant pollution axes (SS_{RES}).

PC1.500

Model	SS_{RES}
BVSTEP subset (16 vars), BC, sqrt	7.224
Sensitivity subset (22 vars), BC, sqrt	7.582
DISTLM BIC subset (7 vars), Euc, ln(x+1)	8.786
BC, sqrt	9.252
Modified Gower	9.540
BVSTEP subset (17 vars), Euc, ln(x+1)	9.718
Sensitivity subset (22 vars), Euc, ln(x+1)	10.657
Euc, ln(x+1)	13.436
Ecological subset (43 vars), BC, sqrt	56.353

PC1.63

Model	SS_{RES}
Modified Gower	8.318
BC, sqrt	8.722
Sensitivity subset (22 vars), BC, sqrt	10.217
Sensitivity subset (22 vars), Euc, ln(x+1)	11.777
Euc, ln(x+1)	12.253
BVSTEP subset (20 vars), BC, sqrt	13.345
DISTLM BIC subset (8 vars), Euc, ln(x+1)	13.548
Ecological subset (43 vars), BC, sqrt	16.546
BVSTEP subset (12 vars), Euc, ln(x+1)	34.175

Split analyses (PC1.63 for group C and PC1.500 for group F)

Model	SS_{RES}
BC, sqrt	11.575
BVSTEP subset[1], BC, sqrt	11.938
DISTLM BIC subset[2], Euc, ln(x+1)	16.966

[1](22 vars for C and 16 vars for F)

[2](28 vars for C and 6 vars for F)

The best model for PC1.500 that actually used all of the biological data (102 taxa) was that obtained using Bray-Curtis on square-root transformed abundances. This CAP model analysis itself, with validation sites superimposed, is shown in Fig. 18(a). Fig. 18(b) shows a scatter plot of predicted versus actual values of the validation sites along PC1.500 for this model. The line with a slope of 1 and intercept of 0 is drawn to help interpret the positions of the points. If prediction is exact, the points would lie precisely on this line. In this model, for example, site 44 is above the line and thus was predicted from the biota to be more polluted than it actually was (i.e., to contain greater total

metal concentrations than it actually did). In contrast, site 21 was predicted to be less polluted than it actually was. This latter kind of error is more dangerous and has more dire consequences, from a management perspective. Thus, models that minimize errors in this direction should be preferred. Also shown on the plot is *b*, the slope of the relationship and *r*, the strength of the relationship (correlation) between the predicted and the actual values. A good model will have both its *b* and *r* values as close as possible to 1 in value.

To understand how the model validation procedure works, let's take an example and trace it through the process. Consider site 18 (Herald Island). Using all biotic variables, we start by calculating the Bray-Curtis dissimilarity between site 18 and each of the 81 samples in the model set (after sqrt transformation) and, based on these dissimilarities alone, place it in the multivariate space and thus project its position along the CAP axis (the *x*-axis of Fig. 18(a)). In this case, the CAP axis value happens to be 0.0642. Now, using the model slope (*b* = 12.12) and intercept (*a* = 0.00), we can therefore calculate the predicted position of this point along the pollution gradient PC1.500, namely:

PC1.500 = 0.00 + 12.12 × CAP = 12.12 × 0.0642 = 0.778. Once we have obtained this predicted value (0.778), we can then compare it with the actual value along PC1.500 obtained according to the metal concentrations measured at the site and the equation given in section 4.1. Plugging in the values of x_{Cu} = 7.73, x_{Pb} = 15.067 and x_{Zn} = 75.33, we obtain an actual value along PC1.500 for site 18 of –0.438 (as shown in Table 20). The predicted value of 0.778 is clearly well above the actual value of –0.438, indicating that the model, in this case, overestimated the degree of pollution at the site. This is reflected in the plot in Fig. 18(b), which shows point 18 to be quite some distance above the line of perfect prediction (with a slope of 1.0).

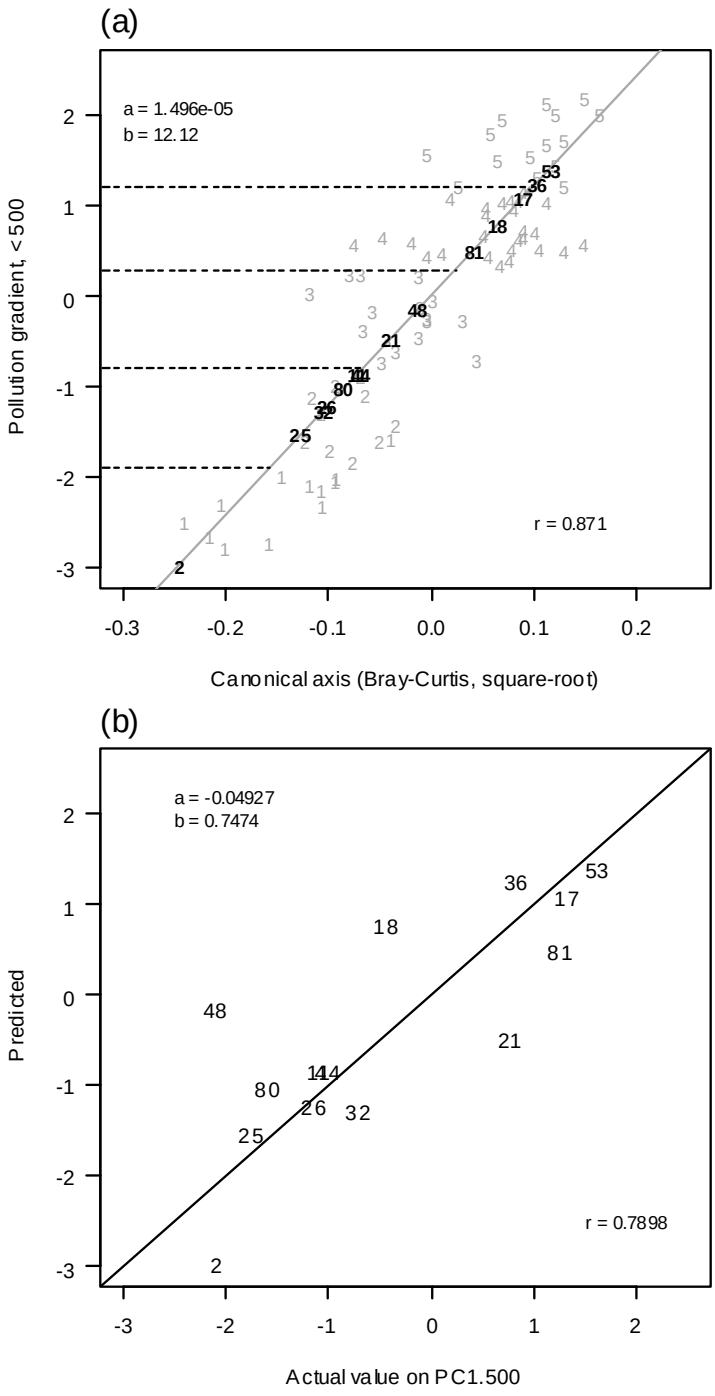


Fig. 18. (a) CAP model of pollution gradient for the whole sample (< 500 μm, top) on the basis of Bray-Curtis dissimilarities of square-root transformed abundances, with validation sites shown in black and (b) predicted values *vs*. actual values for validation sites along PC1.500 for the model. The line on the plot has a slope of 1.0 and all points would lie on the line if prediction from the model were perfect.

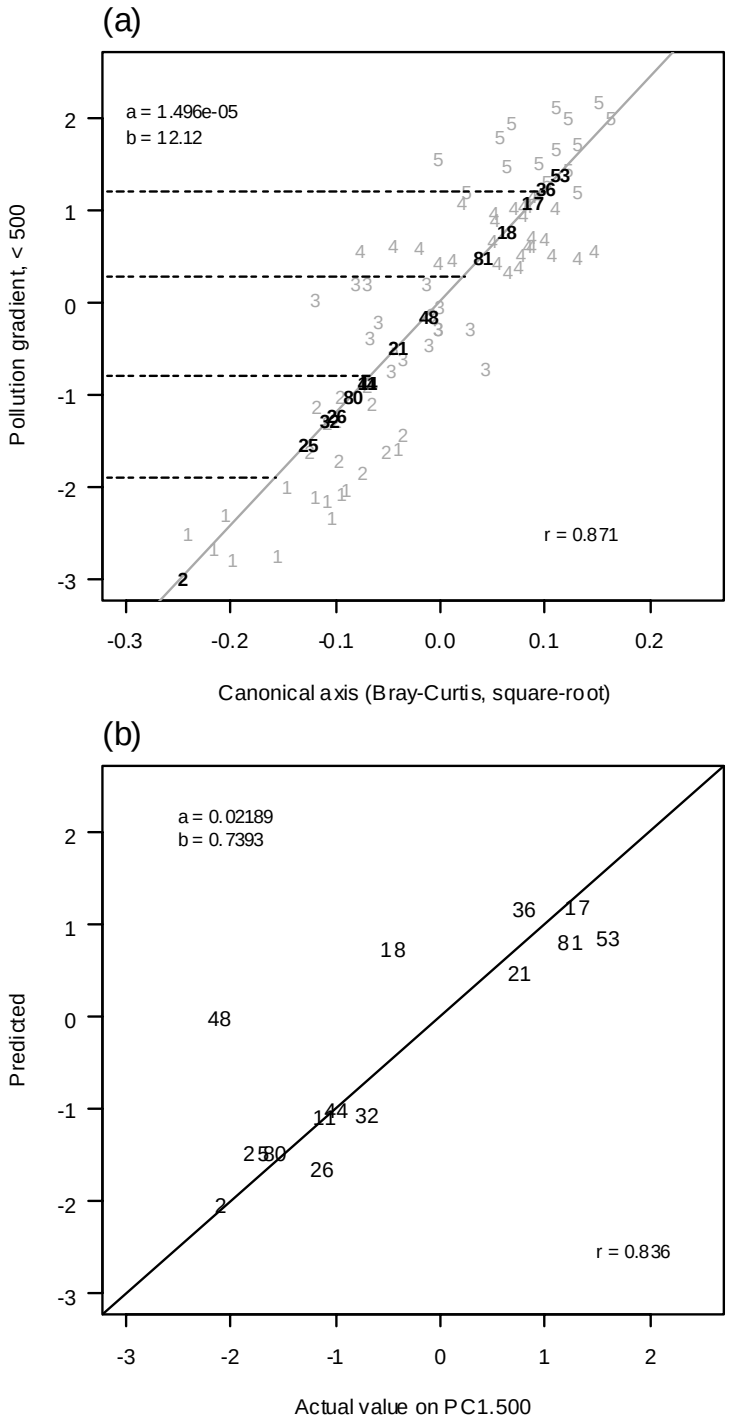


Fig. 19. (a) CAP model of pollution gradient for the whole sample (< 500 μm, top) on the basis of Bray-Curtis dissimilarities of square-root transformed abundances, with validation sites shown in black, using only a subset of 16 variables identified from BVSTEP and (b) predicted values *vs.* actual values for validation sites along PC1.500 for the model. The line on the plot has a slope of 1.0 and all points would lie on the line if prediction from the model were perfect.

In general, however, the benthic health model obtained using CAP on Bray-Curtis using all the data appears to do a pretty reasonable job of predicting the positions of validation sites. Most of the validation sites lie close to the line and the correlation between

predicted and actual values is 0.79. Nevertheless, considerable improvement is obtained using the best overall model: the BVSTEP subset (16 taxa) with Bray-Curtis on square-root transformed data (Fig. 19). Here, almost all of the points lie close to the line and the strength of the relationship between predicted and actual values is improved (r = 0.84). What is more, those sites which are deviating from the line (i.e. site 48 and site 18) do so in a way that would lead to a conservative remedial action in line with the precautionary principle. That is, the model predicts that these particular sites are more polluted than they actually are. Although it cannot be stated that this pattern of conservatism would necessarily be repeated with a new set of validation sites, the combination of accuracy and conservatism seen here does bode well for the use of this model in general.

The subsets obtained using either the sensitivity subset or the DISTLM subset did not appear to be as successful as the BVSTEP subset, at least for these validation sites (Fig. 20). Although the degree of correlation (r) was comparable to that obtained by the BVSTEP subset, a fairly important difference was seen in the size of the slope (b). For these two subsets, the slope of the relationship between predicted and actual values moved closer to zero. Thus, even though the degree of relationship was fairly strong, the decrease in the value of the slope away from a value of 1 and towards zero indicated that the replicability of the pollution gradient (as in a 1:1 relationship) was not as good for these models. These models would have a tendency to predict healthy sites to be more polluted than they actually are and to predict polluted sites to be healthier than they actually are.

For PC1.63, the best model was obtained using the Modified Gower measure on all taxa in the biotic community (Table 21, Fig. 21). This model performed quite well for the majority of the validation sites: many of the predictions are extremely close to their actual values along the PC1.63 axis (Fig. 21(b)). It appears that the model essentially had trouble placing sites 48, 44 and 11 onto their correct positions along the axis (Fig. 21(b)). Site 48 was overestimated, while sites 44 and 11 were underestimated. The model obtained using Bray-Curtis on square-root transformed data performed similarly to the model based on the Modified Gower measure; it also apparently had trouble placing these three points, in particular, onto their correct positions on the axis (Table 21, Fig. 22(a)). Unfortunately, classification of the sites into groups based on their physical variables and the use of separate models for these (groups C and F) resulted in no clear improvement in terms of predictive capability (Fig. 22(b), (c)). This was rather surprising, given the high canonical correlation for PC1.63 versus group C (Table 17, Fig. 14). As stated earlier, however, a high canonical correlation does not necessarily mean good predictive power. For example, one can achieve a high canonical correlation simply by increasing the number of PCO axes (m) to use in the CAP analysis. This is why the leave-one-out residual sum of squares (or allocation success, in the case of groups) is an essential statistic to use in the interpretation of the potential utility of a CAP model. It also emphasizes that no model's worth can be understood without some independent model validation procedure, as has been done here.

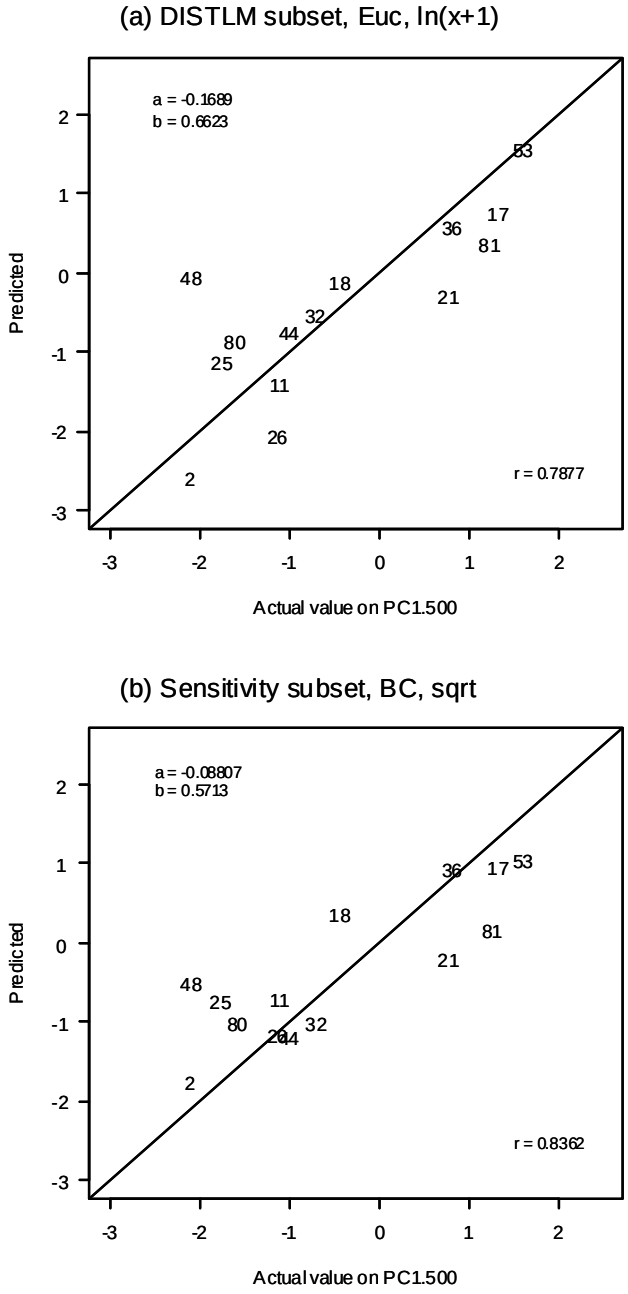


Fig. 20. Predicted values *vs.* actual values along PC1.500 for each of two CAP models: (a) based on the subset of taxa obtained using the BIC criterion and DISTLM, with Euclidean distances on ln(x+1)-transformed abundances and (b) based on the *a priori* chosen sensitivity subset with Bray-Curtis dissimilarities on square-root transformed abundances. The line on each plot has a slope of 1.0 and all points would lie on the line if prediction from the model were perfect.

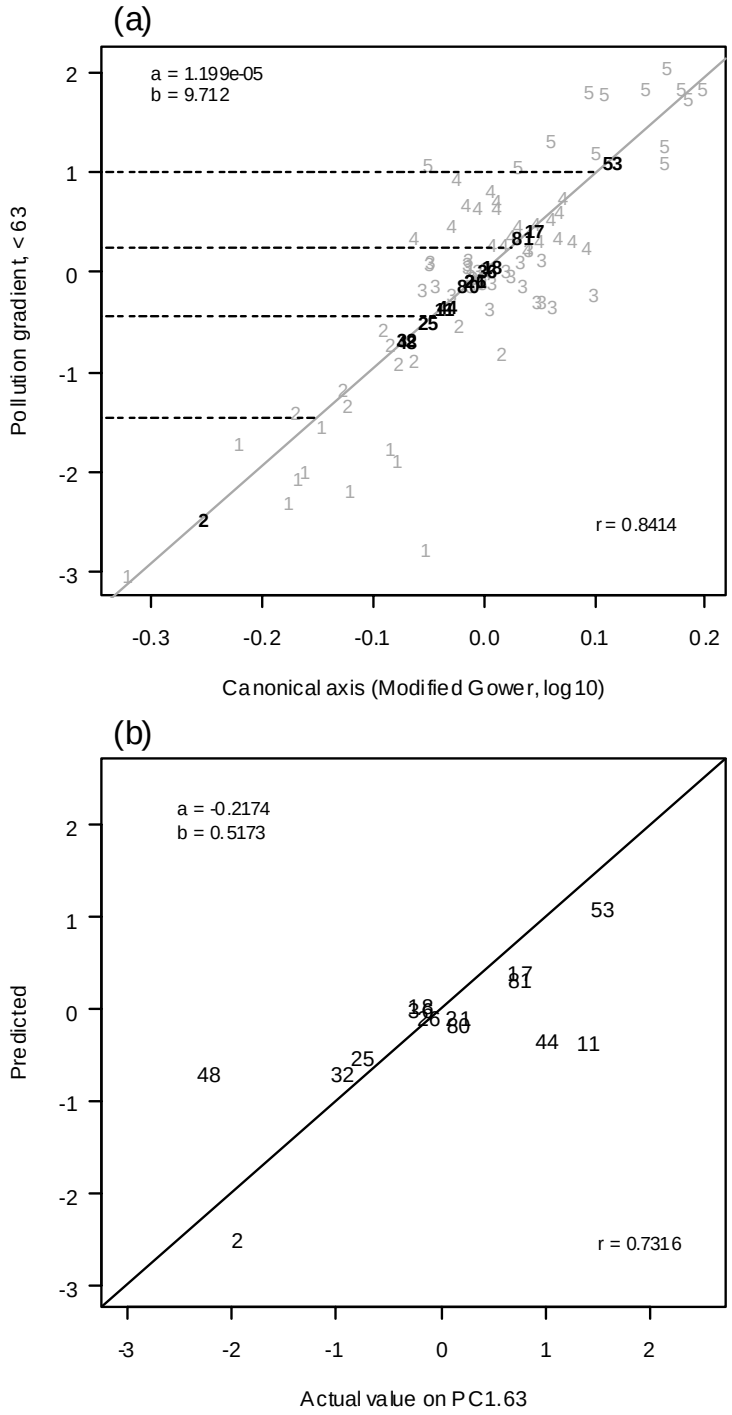


Fig. 21. (a) CAP model of pollution gradient for the mud fraction (< 63 μm, top) on the basis of Modified Gower dissimilarities (log base 10), with validation sites and (b) predicted values *vs.* actual values along PC1.63 for the model. The line on the plot has a slope of 1.0 and all points would lie on the line if prediction from the model were perfect.

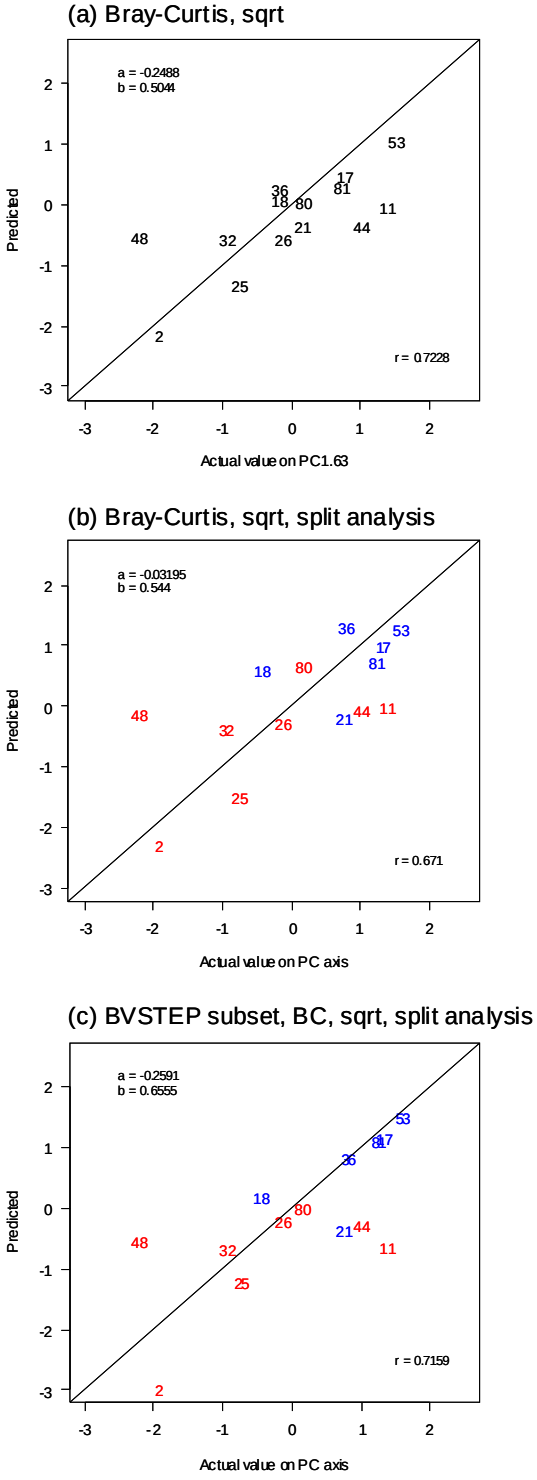


Fig. 22. Predicted *vs.* actual values for each of three CAP models: (a) PC1.63 modelled by all taxa on the basis of Bray-Curtis, square-root transformed abundances, (b) a split model based on all taxa and (c) a split model based on BVSTEP subsets of taxa. For split models, separate CAPmodels were applied to samples in group C (red) for PC1.63 and in group F (in blue) for PC1.500.

7 DISCUSSION

7.1 Relating PC gradients to existing sediment quality guidelines

High correlation among the three metal contaminants, copper (Cu), lead (Pb) and zinc (Zn), measured from sites across the region indicated that elevated levels of these metals currently tend to occur together in soft-sediment intertidal benthic habitats. Potential ecological impacts of simultaneous multiple stressors are not well understood and are the subject of ongoing research. However, this high degree of spatial correlation among the metals allowed a single pollution gradient to be derived from the data, using principal components analysis (PCA). A separate gradient was built from measurements of metal concentrations from the total sediment sample (PC1.500) and from the mud fraction (PC1.63), and 5 groups of samples along each gradient were identified using a *k*-means partitioning algorithm. Each PC axis is a linear combination of these three metal variables, so new samples can be positioned along each of these axes, given measured concentrations of each of the three metals.

There are a number of existing guidelines from various sources which provide threshold values for these metals in sediments, above which a given site might be classified as "polluted" (Table 22). Using some of these guidelines, the ARC has developed and adopted certain "Environmental Response Criteria" (ERC) for assessing the level of pollution in Auckland's benthic intertidal habitats (ARC 2004). These criteria can be used to classify sites into one of three categories: "green" (= healthy), "amber" or "red" (= polluted) on the basis of the levels of individual contaminants (Table 22).

Table 22. Existing sediment quality guidelines from various sources, as indicated, along with their positions along PC axes derived in the present study. ISQG = Interim Sediment Quality Guidelines and ERC = Environmental Response Criteria. Values for the metals are given as concentrations in mg/kg.

Source	Guideline	Zn	Cu	Pb	PC1.500	PC1.63
ANZECC (2000)	ISQG-Low	200	65	50	2.090	1.418
ANZECC (2000)	ISQG-High	410	270	220	4.213	3.530
Long et al. (1995)	Effects-Range Low (ER-L)	150	34	46.7	1.500	0.880
Long et al. (1995)	Effects-Range Median (ER-M)	410	270	218	4.208	3.524
MacDonald (1996)	Threshold Effects Level (TEL)	124	18.7	30.2	0.776	0.176
MacDonald (1996)	Probable Effects Level (PEL)	271	108	112	3.035	2.381
ARC Blueprint (2004)	ERC-Green	<124	<19	<30	0.782	0.180
ARC Blueprint (2004)	ERC-Amber	124-150	19-34	30-50		
ARC Blueprint (2004)	ERC-Red	>150	>34	>50	1.540	0.925

By simply "plugging in" the values for the metal concentrations provided by these guidelines into the equation given in section 3.1 above, we can determine the value they obtain along each of the PC axes we have derived, providing a clearer context for

the present study and its potential value for ongoing monitoring in the region (Figs. 23, 24).

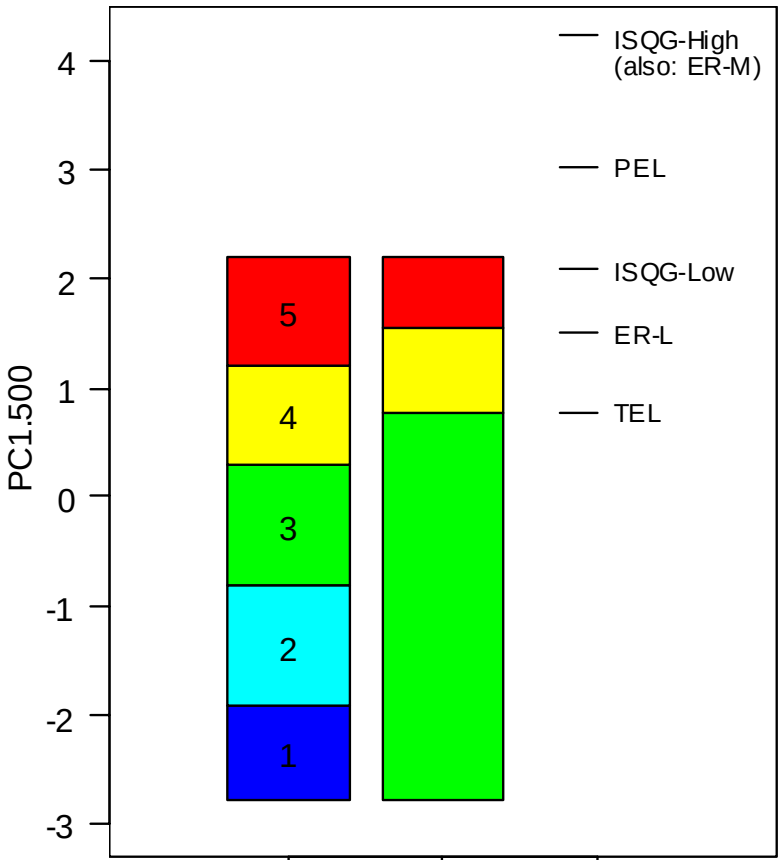


Fig. 23. Values of sediment quality guidelines given in Table 22 along the pollution gradient for the total sediment sample, PC1.500. The first bar shows the 5 groups derived in the present study, next is shown the environmental response criteria (ERC) as green, amber and red. Other guidelines are shown as single values along the axis.

From these graphics, it is clear that ISQG-High, ER-M and PEL indicate extremely high levels of pollution that are all out of range of the measures we have from sites in our study of the Auckland Region. Also, for PC1.500, our groups labeled 4 and 5 occur at slightly lower values along the pollution gradient than do the existing ERC amber and red guidelines (Fig. 23). Our groupings are also more refined over the whole range of sites and yield discrimination along the lower end of the scale, suggesting that community health is affected below the current guidelines, and providing earlier warning signs of pollution. Similar observations may be made for PC1.63, although in this case, our groups 4 and 5 correspond well with the existing amber and red guidelines of the ERC. As for PC1.500, this model also provides better discrimination for variation in lower values (1, 2 and 3) along the pollution gradient (Fig. 24).

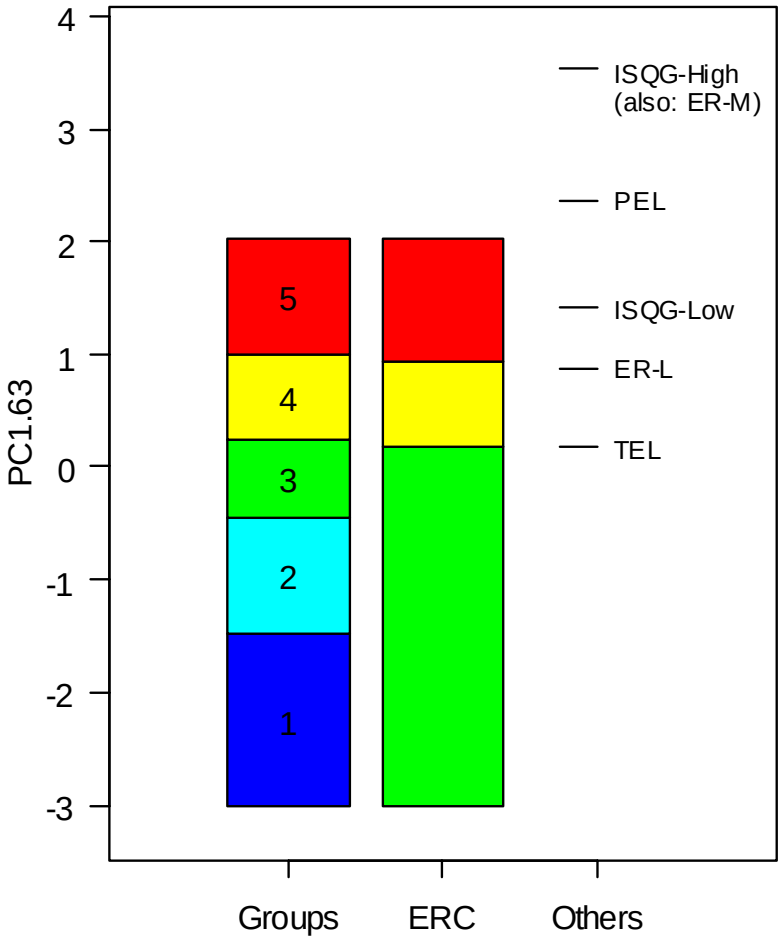


Fig. 24. Values of sediment quality guidelines given in Table 22 along the pollution gradient for the mud fraction of a sample, PC1.63. The first bar shows the 5 groups derived in the present study, next is shown the environmental response criteria (ERC) as green, amber and red. Other guidelines are shown as single values along the axis.

These results suggest that the current criteria are too high. Sites occurring in group 5 are polluted, according to current criteria and the health model. However, rather than considering those in group 4 as borderline, sites in group 3 may be considered as a more "critical" group. Sites in this category should perhaps be receiving the greatest attention with respect to both monitoring and potential remedial management action. Sites in groups 1 or 2 can be considered as healthy, although changes in community structure are beginning to be detected. This higher sensitivity may be a result of: (i) differences between field situations and laboratory tests, or (ii) differences between guidelines based on single contaminants and those measuring multiple stresses.

Some caveats are important to keep in mind with respect to the proposed PC pollution axes. First, their use relies upon there being high correlation in the levels of copper, zinc and lead in the environment across the region. This is appropriate in our current context; the PC axes will give a clear signal that integrates information about these

simultaneous stressors genuinely occurring together in the environment. However, if a site were to experience increases in just one or two of the three contaminants, but not in all of them, then the sediment quality guidelines which provide individual threshold values for each metal would be more sensitive than the overall pollution gradient measure at picking up such a change. So, our derived PC axes should be used in conjunction with existing individual metal-based criteria.

Although concentrations of the three metals are highly correlated, zinc levels tend to be closer to upper guideline values and therefore have the potential to have a greater influence on biological effects than the other two contaminants. In addition, lead levels are more variable through space and time, due to the removal of the key source of lead (i.e., leaded petrol). Lead levels are therefore likely to decline through time, whereas copper and zinc levels are likely to continue to increase (Williamson and Mills 2002). Relationships between existing concentrations and accumulation rates demonstrate this (S. Kelly, pers.comm.). Thus, although the pollution gradient identified here is likely to prove very useful in the immediate term, the separate guidelines for individual metal concentrations should not be abandoned. Longer term changes in relative metal concentrations across the region may also result in the evolution of new models in the future that may treat metals separately, especially if the degree of spatial correlation among metals in the environment deteriorates.

7.2 The best models of benthic ecosystem health

Ecological assemblages generally reflected pollution gradients very well, all along their range. The present study identified clear methods for modeling the pollution gradient axes using ecological data. The best models of benthic ecosystem health were those which obtained high canonical correlations with the pollution gradient(s) and which had a low level of error when new sites were tested (validation).

The ARC's "Blueprint for monitoring urban receiving environments" (2004) indicated that the ERC guidelines should be applied using metal concentrations from the total sediment sample (<500 μm) for sites in the inner Settling Zones and using metal concentrations from the mud fraction (< 63 μm) for sites in the Outer Zones. Here, we used the actual characteristics of the sediment (grain size fractions) from each site as well as explicit exposure indices (furthest and closest wind exposure) to identify two physical groups of sites: those having coarser sediments and greater exposure (group C) and those having finer sediments and lesser exposure (group F).

The best overall ecological models were obtained using all sites together, rather than splitting them into two groups, and the biotic assemblages had the strongest relationships with metal concentrations in the total sediment sample (PC1.500). This supports the notion that the biota are responding to all metals in the sediments, and not just to those considered to be bioavailable in the mud fraction. Although group C and group F corresponded roughly with the existing designations of Outer Zones and

Settling Zones, respectively, better models were obtained using the site designations of C and F. No advantage was obtained by relating sites in group F alone to total metal concentrations. However, excellent models were obtained by considering sites in group C alone and relating these to metal concentrations in the mud fraction (PC1.63, δ = 0.985, Bray-Curtis, sqrt). These results therefore also support the notion that heavy metals are potentially more bioavailable in the mud fraction in outer zones (Williamson and Mills 2002).

We consider that the following models qualify as "best" models from the work done here.

For all sites, regardless of their physical characteristics:

- PC1.500 vs. Bray-Curtis dissimilarities on square-root transformed abundances of all taxa in the assemblage (canonical correlation δ = 0.871).
- PC1.500 vs. Bray-Curtis dissimilarities on square-root transformed abundances of the BVSTEP subset of 16 taxa (δ = 0.865).
- PC1.500 vs. Bray-Curtis dissimilarities on square-root transformed abundances of the sensitivity subset of 22 taxa (δ = 0.808).

Also, for sites having physical characteristics of coarser sediments and greater exposure (i.e., sites in group C only), an excellent model was obtained for metal concentrations in the mud fraction:

- PC1.63 vs. Bray-Curtis dissimilarities on square-root transformed abundances of all taxa *for sites in group C* (δ = 0.985).

Although not as strong as models using metals from the total sample, the best overall models using metal concentrations in the mud fraction were:

- PC1.63 vs. Modified Gower dissimilarities (log base 10) of all taxa (δ = 0.841).
- PC1.63 vs. Bray-Curtis dissimilarities on square-root transformed abundances of all taxa (δ = 0.787).

7.3 Some comments on models using subsets of taxa

Models using subsets of taxa (the BVSTEP subset with 16 variables or the sensitivity subset with 22 variables) gave surprisingly good results for PC1.500 when it came to model validation using new sites. It is tempting to consider using only a subset of taxa for benthic health models, as sampling subsets might be more efficient than enumerating and identifying all of the taxa present at a site. There are, however, some potential pitfalls with using subsets of taxa for assessing ecosystem health.

First, subsets chosen using statistical methods may have some practical problems under longer-term scrutiny. First, the subset identified by BVSTEP included some fairly rare taxa (e.g., *Trochodota dendyi*, *Haminoea zelandiae*, *Cirolana sp.*, *Edwardsia sp.*, see

frequencies of occurrence in Appendix 2). The inherent normalization of variables in the CAP analysis may result in there being an emphasis on such rarer taxa if their occurrence happens to discriminate sites along the pollution gradient well. However, patchily distributed species tend to have highly variable distributions and thus may not actually provide good predictive properties for longer term modeling.

In addition, some of the taxa in subsets may have been chosen because they act as proxies for other taxa. For example, *Anthopleura aureoradiata* (chosen in the BVSTEP subset) is an anemone that attaches to hard surfaces and tends to occur on the shells of cockles, *Austrovenus stutchburyi*, in these soft-sediment environments. *Anthopleura* may have been chosen in the subset because of a strong correlation with *Austrovenus* (Spearman's $\rho = 0.668$), whereas the latter might actually respond more sensitively to metal contamination. It may not matter at present that correlations among species induce certain proxy organisms to be chosen in subsets, but if pollution were to alter the nature of these biological associations (e.g., perhaps *Anthopleura* in time could occur on dead shells of *Austrovenus* just as readily as on live shells), then the utility of the subset would lose its force.

Thus, if a subset were to be used, then one which has been motivated from biological knowledge may be more defensible in the long run than one which was obtained *via* statistical methods. Although the sensitivity subset is certainly a candidate in this respect (and it seemed to perform reasonably well for validation), the canonical correlation for the model using this subset ($\delta = 0.808$) was not as high as that obtained when all taxa were used ($\delta = 0.871$). This suggests that, at present, the CAP model of PC1.500 versus all taxa (BC, sqrt) should be the one used routinely for monitoring and management. Models using subsets (either the 22-variable sensitivity or the 16-variable BVSTEP subset) can also be used presently with a high degree of confidence in the event that enumeration of all taxa is not possible.

7.4 Recommendations

We recommend the use of the following two models in tandem:

- PC1.500 vs. Bray-Curtis dissimilarities on square-root transformed abundances of all taxa (canonical correlation $\delta = 0.871$).
- PC1.63 vs. Bray-Curtis dissimilarities on square-root transformed abundances of all taxa *for sites in group C only* ($\delta = 0.985$).

An outline of appropriate steps to follow given one or more new samples of ecological data is shown in Fig. 25 and detailed below (numbers for steps here correspond to numbers given in the figure):

1. Bray-Curtis dissimilarities between the new samples and existing samples (square-root transformed) can be used to place the new samples onto the CAP axis in the model of PC1.500 versus all taxa.

2. The values of PC1.500 for the new samples can then be predicted from the values of each of them along the CAP axis.

3. The boundary values given in Table 6 along PC1.500 can be used to classify each new sample into one of the groups from 1 to 5 ("the first classification").

4. If physical data are available, use them to classify the new samples as belonging either in group C or group F. To err on the side of caution, borderline samples should be considered as belonging in group C. If physical data are unavailable, determine whether the site may be considered to belong in the Outer Zone.

5. For sites in group C (or in the Outer Zone), Bray-Curtis dissimilarities between the new samples and existing samples in group C only (square-root transformed) can be used to place the new samples onto the CAP axis in the model of PC1.63 versus all taxa.

6. The values of PC1.63 for the new samples can then be predicted from the values of each of them along this CAP axis.

7. Use the boundary values given in Table 6 along PC1.63 to classify each new sample into one of the groups from 1 to 5 ("the second classification").

8. For each of the new samples, take the larger of the first and second classifications (obtained in steps 3 and 7 above, respectively) as the most cautious estimator of the current state of benthic health (1 = healthy and 5 = polluted). New samples in group F will use only the first classification.

9. If metal concentration data are available, use the direct equations for each of PC1.500 and PC1.63 along with Table 6 in order to validate results obtained using the ecological data. In addition, the ERC guidelines can also be applied on the basis of individual metal concentrations in this case.

By using the above outlined approach, these models can be used to monitor and make management decisions concerning benthic ecosystem health in the Auckland Region. Appropriate software (PRIMER v6, with the add-on PERMANOVA+ , by Anderson and Gorley, to be released in 2007) and guidance on its use will be provided to the ARC in order to perform all necessary calculations and implement these techniques. We also recommend that the degree of correlation among metal contaminants be checked through time to ground-truth the ongoing utility of the PC axes as overall pollution gradients. In addition, if new species or taxa are sampled in future that are not included in the current models, then these should simply be omitted. If, however, the number of new taxa encountered increases and they are sampled consistently, then new models should be developed to include them.

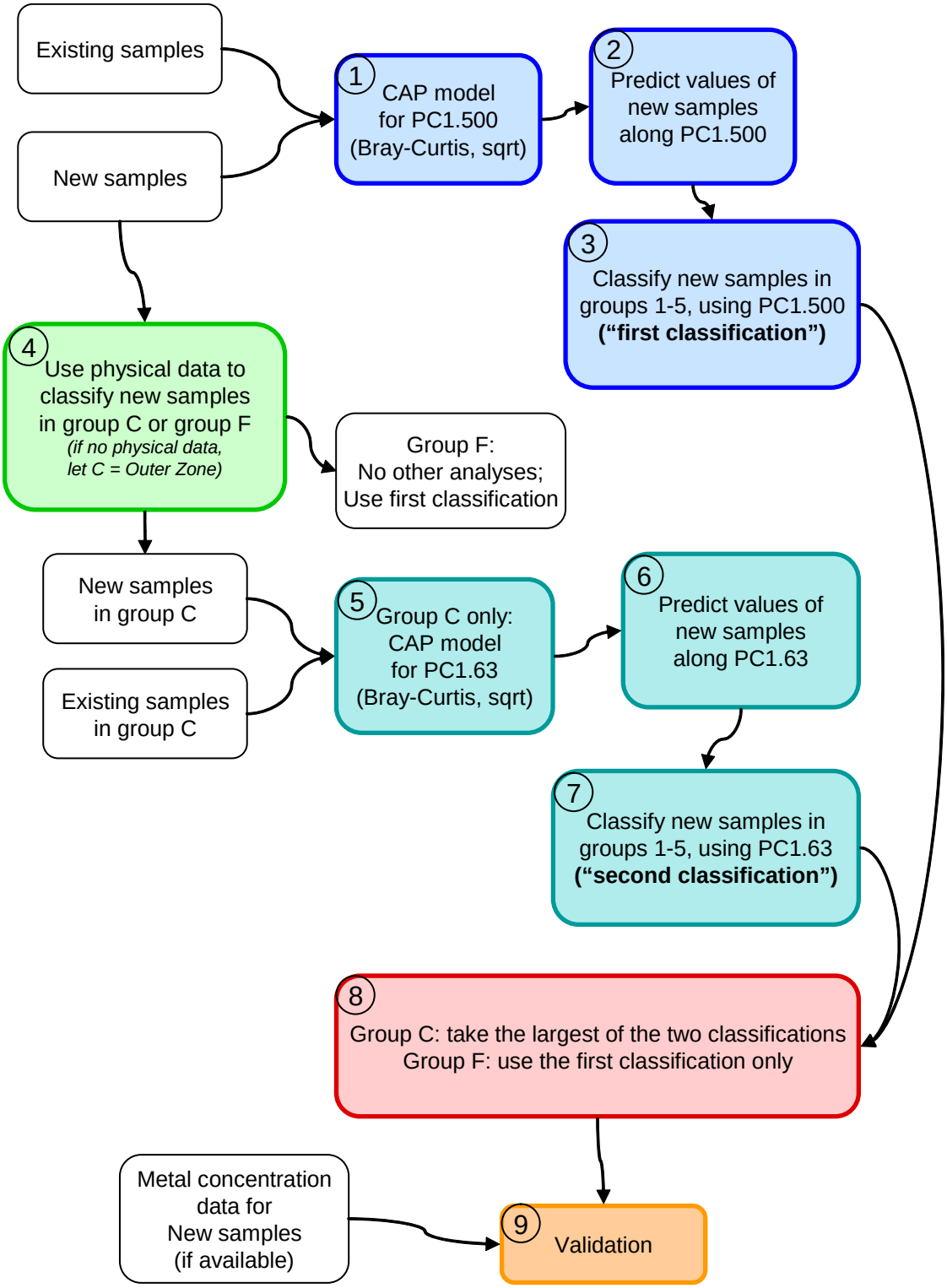


Fig. 25. Flowchart of recommended analyses to assess benthic ecosystem health.

8 REFERENCES

Akaike, H. 1973. Information theory as an extension of the maximum likelihood principle. pp. 267-281 in Petrov, B.N. and Caski, F. (eds.). *Proceedings, 2nd International Symposium on Information Theory*. Budapest, Akademiai Kiado.

Anderson, M.J., Ellingsen, K.E., McArdle, B.H. 2006. Multivariate dispersion as a measure of beta diversity. *Ecology Letters* 9: 683-693.

Anderson, M.J., Hewitt, J.E. and Thrush, S.F. 2002. *Using a multivariate statistical model to define community health*. NIWA Client Report: HAM2002-048, prepared for Auckland Regional Council.

Anderson, M.J. and Robinson, J. 2003. Generalized discriminant analysis based on distances. *Australian & New Zealand Journal of Statistics* 45: 301-318.

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

ANZECC 2000. *Australian and New Zealand guidelines for fresh and marine water quality. National water quality management strategy*. Australian and New Zealand Environment and Conservation Council. Agriculture and Resource Management Councils of Australia and New Zealand. Canberra, Australia.

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

Auckland Regional Council. 2004. *Blueprint for monitoring urban receiving environments*. TP168. 66 pp.

Box, G.E.P. and Cox, D.R. 1964. An analysis of transformations. *Journal of the Royal Statistical Society, Series B*, 26: 211–252.

Calinski, T. and Harabasz, J. 1974. A dendrite method for cluster analysis. *Communications in Statistics* 3: 1-27.

Chao, A., Chazdon, R.L., Colwell, R.K. and Shen, T. 2005. A statistical approach for assessing similarity of species composition with incidence and abundance data. *Ecology Letters* 8: 148-159.

Clarke, K.R. and Gorley, R.N. 2006. *PRIMER v6: User Manual/Tutorial*. PRIMER-E: Plymouth, UK.

Clarke, K.R., Somerfield, P.J. and Chapman, M.G. 2006. On resemblance measures for ecological studies, including taxonomic dissimilarities and a zero-adjusted Bray-Curtis coefficient for denuded assemblages. *Journal of Experimental Marine Biology and Ecology* 330: 55-80.

Ford, R.B. and Anderson, M.J. 2005. *Ecological monitoring of the Okura and Whitford estuaries 2004-2005: temporal and spatial extensions of regional models*. Auckland UniServices Report for the Auckland Regional Council, TP 287, Auckland, 83 pp.

Funnell, G.A., Hewitt, J.E. and Thrush, S.F. 2003. *Ecological monitoring programme for Manukau Harbour: report on data collected up to February 2003*. Auckland Regional Council, TP 224, Auckland, 36 pp.

Gower, J.C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* 53: 325-338.

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

Hewitt, J.E., Lundquist, C., Hancock, N., Halliday, J., Chiaroni, L. 2004. *Waitemata Harbour ecological monitoring programme: summary of data collected from October 2000-February 2004*. Auckland Regional Council, Auckland.

Hewitt, J.E., Lundquist, C., Halliday, J. and Hickey, C. 2006. *Upper Waitemata Harbour ecological monitoring programme – results from the first year of monitoring 2005-2006*. Auckland Regional Council, Auckland 65 pp.

Kingett Mitchell Limited. 2002. *Auckland regional benthic ecology monitoring 2002*. Unpublished report for Auckland Regional Council.

Lachenbruch, P.A. and Mickey, M.R. 1968. Estimation of error rates in discriminant analysis. *Technometrics* 10: 1-11.

Legendre, P. and Legendre, L. 1998. *Numerical ecology, 2nd English edition*. Elsevier, Amsterdam.

Long, E.R.m MacDonald, D.D., Smith, S.L. and Calder, F.D. 1995. Incidence of adverse biological effects within ranges of chemical concentrations in marine and estuarine sediments. *Environmental Management* 19:81-97

MacDonald, D.D., Carr, R.S., Calder, F.D., Long, E.R. and Ingersoll, C.G. 1996. Development and evaluation of sediment quality guidelines for Florida coastal waters. *Ecotoxicology* 5: 253-278.

MacQueen, J. 1967. Some methods for classification and analysis of multivariate observations. Pp. 281-297 in: Le Cam, L.M. and Neyman, J. (eds.) *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability. Vol. 1*. University of California Press, Berkeley.

McArdle, B.H. and Anderson, M.J. 2001. Fitting multivariate models to community data: a comment on distance-based redundancy analysis. *Ecology* 82: 290-297.

Millar, R.B., Anderson, M.J. and Zunun, G. 2005. Fitting nonlinear environmental gradients to community data: a general distance-based approach. *Ecology* 86: 2245-2251.

Potter, I.C., Bird, D.J., Claridge, P.N., Clarke, K.R., Hyndes, G.A., Newton, L.C. 2001. Fish fauna of the Severn Estuary. Are there long-term changes in abundance and species composition and are the recruitment patterns of the main marine species correlated? *Journal of Experimental Marine Biology and Ecology* 258: 15-37.

R Development Core Team. 2005. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

Reed, J. and Webster, K. 2004. *Marine sediment monitoring programme: 2003 results*. Auckland Regional Council, TP 246, Auckland, 146 pp.

Schwarz, G. 1978. Estimating the dimension of a model. *Annals of Statistics* 6: 461-464.

Seber, G.A.F. and Lee, A.J. 2003. *Linear regression analysis, 2nd edition*. John Wiley & Sons, Hoboken, New Jersey.

Venables, W.N. and Ripley, B.D. 2002. *Modern applied statistics with S. 4th edition*. Springer-Verlag, New York.

Warton, D.I. 2005. Many zeros does not mean zero inflation: comparing the goodness-of-fit of parametric models to multivariate abundance data. *Environmetrics* 16(3): 275-289.

Williamson, R.B. and Kelly, S. 2003. *Regional discharges project marine receiving environment status report 2003*. Auckland Regional Council, TP 203, Auckland, 53 pp.

Williamson, R.B. and Mills, G.N. 2002. *Sediment quality guidelines for the Regional discharges project*. Diffuse Sources Ltd Report for the Auckland Regional Council, Auckland, 61 pp.

9 Appendix 1. List of samples used for modeling and validation.

There were 95 samples in total obtained from 84 sites across the region. Each sample is uniquely identifiable for a given site in a give year. Coordinates for each site are given according to the New Zealand Map Grid. The column "Mod.Val" indicates whether a sample was used for modeling (M) or for validation (V). An asterisk indicates that sediment texture for that site was actually obtained in May 2006, even though the biology was sampled in the year shown (see text for details). The column headed "Source" identifies the source of data, with the first integer indicating the source for sediment data, and the second integer indicating the source for biological data, as follows: 1 = Kingett Mitchell Limited (2002), 2 = Funnell et al. (2003), 3 = Hewitt et al. (2004), 4 = Reed and Webster (2004), 5 = Williamson and Kelly (2003), 6 = new unpublished data collected by the ARC, 7 = Ford and Anderson (2005), 8 = Hewitt et al. (2006).

Site no.	Site name	Easting	Northing	Mod.Val	Year	Source
1	Anns Creek	2672634	6473059	M	2002	4, 1
1	Anns Creek	2672634	6473059	M	2005	6, 6
2	Auckland Airport	2672515	6463388	V	2002	5, 2
3	Awatea Rd	2670466	6481358	M	2004	6, 6
4	Bengazi	2677243	6476984	M	2004	6, 6
5	Bowden Rd	2675670	6474640	M	2004	6, 6
6	Brigham	2653704	6490358	M	2005	8, 8
7	Cape Horn	2659917	6470448	M	2002	5, 2
8	Chelsea	2664602	6485392	M	2004	6, 6
9	Clarkes Beach	2661675	6452219	M	2002	5, 2
10	Coxes, Waitemata	2663914	6482238	M	2004	6, 6
11	Coxs	2664141	6482090	V	2005*	6, 6
12	Glendowie	2678366	6479045	M	2005*	6, 6
13	Hellyers	2661837	6489965	M	2005*	6, 6
14	Hellyers outer	2660692	6489573	M	2005	8, 8
15	Henderson Entrance	2658591	6486244	M	2002	5, 3
15	Henderson Entrance	2658591	6486244	M	2004	6, 6
16	Henderson Lower	2656715	6484645	M	2004	6, 6
17	Henderson Upper	2656017	6483479	M	2002	5, 1
17	Henderson Upper	2656017	6483479	V	2005	6, 6
18	Herald Island	2658153	6489728	V	2005*	6, 6
19	Hi North	2658478	6490325	M	2005	8, 8
20	Hillsborough	2667210	6473280	M	2004	6, 6
21	Hobson - Purewa Bridge	2672100	6480461	V	2005*	6, 6
22	Hobson - Tohunga	2670174	6480830	M	2005*	6, 6
23	Hobsonville	2660106	6487972	M	2002	5, 3
23	Hobsonville	2660106	6487972	M	2005	6, 6
24	Kaipatiki	2661944	6489895	M	2005	6, 6

Site no.	Site name	Easting	Northing	Mod.Val	Year	Source
25	Kendalls	2662758	6484849	V	2004	6, 6
26	Little Shoal Bay	2665796	6485551	V	2005*	6, 6
27	Lower Shoal Bay	2667976	6486007	M	2005	6, 6
28	Lucus outer	2658788	6491194	M	2005	8, 8
29	Lucus Te Wharau	2659685	6491950	M	2004	6, 6
30	Lucus Upper	2660154	6492967	M	2005	6, 6
31	Mangemangeroa B	2684748	6474915	M	2004*	6, 6
32	Mangemangeroa E	2684488	6474876	V	2004*	6, 6
33	Mangere Cemetery	2670400	6472900	M	2005	6, 6
34	Mangere Inlet: Harania Creek	2672344	6471451	M	2005*	6, 6
35	Mangere Inlet: Kiwi Esplanade	2668945	6472230	M	2005*	6, 6
36	Mangere Inlet: Tararata Creek	2670928	6471434	V	2005*	6, 6
37	Meola Inner	2662817	6481374	M	2002	4, 1
37	Meola Inner	2662817	6481374	M	2005	6, 6
38	Meola Outer	2662751	6481994	M	2004	6, 6
39	Meola Reef	2662897	6482580	M	2002	4, 1
39	Meola Reef	2662897	6482580	M	2005	6, 6
40	Meola West	2662746	6481631	M	2005*	6, 6
41	Middlemore	2675627	6470765	M	2002	4, 1
41	Middlemore	2675627	6470765	M	2005	6, 6
42	Motions	2663020	6481413	M	2002	4, 1
42	Motions	2663020	6481413	M	2005	6, 6
43	Motions East	2663127	6481978	M	2005*	6, 6
44	Newmarket	2670161	6480662	V	2005	6, 6
45	Ngataringa Bay	2670233	6485348	M	2005*	6, 6
46	Oakley	2661590	6479618	M	2005	6, 6
47	Okura D	2664903	6502243	M	2004*	6, 7
48	Okura J	2663712	6501553	V	2004*	6, 7
49	Orewa F	2661291	6509855	M	2004*	6, 7
50	Orewa G	2661690	6509891	M	2004*	6, 7
51	Otahuhu Creek	2675971	6472757	M	2004	6, 6
52	Out Main UWH	2659043	6490098	M	2005	8, 8
53	Pakuranga	2678591	6473361	V	2005	6, 6
54	Pakuranga mid	2677797	6473177	M	2005*	6, 6
55	Panmure	2674935	6475550	M	2004	5, 1
56	Paremoremo	2656364	6492284	M	2005	6, 6
57	Paremoremo upper	2656207	6492093	M	2005	6, 6
58	Pollen Island	2660495	6479910	M	2005	6, 6
59	Princess St	2676230	6472210	M	2004	6, 6
60	Puhinui	2675443	6462162	M	2005*	6, 6
61	Puhinui, Entrance	2675350	6461350	M	2002	5, 1
62	Puhoi F	2663158	6517525	M	2004*	6, 7
63	Puhoi H	2662611	6517889	M	2004*	6, 7

Site no.	Site name	Easting	Northing	Mod.Val	Year	Source
64	Pukaki	2671080	6465235	M	2005*	6, 6
65	Purewa	2672906	6480183	M	2004	6, 6
66	Rangitopuni	2653196	6492099	M	2005*	6, 6
67	Rangitopuni UWH	2653449	6491807	M	2005	8, 8
68	Shoal Bay, Hillcrest	2667825	6487365	M	2004	6, 6
69	Shoal Bay, Upper	2668358	6486410	M	2004	6, 6
70	Turanga G	2685520	6473527	M	2004*	6, 7
71	Turanga J	2685936	6472558	M	2004*	6, 7
72	Upper main UWH	2654360	6491000	M	2005	8, 8
73	Victoria Ave	2671269	6480637	M	2004	6, 6
74	Waiwera E	2663038	6516553	M	2004*	6, 7
75	Waiwera J	2662264	6516484	M	2004*	6, 7
76	Weiti	2662420	6508229	M	2005*	6, 6
77	Whakataka	2671684	6481227	M	2002	5, 1
77	Whakataka	2671684	6481227	M	2005	6, 6
78	Whau East	2659578	6477506	M	2005*	6, 6
79	Whau Entrance	2658515	6482039	M	2004	6, 6
80	Whau Entrance, WHO A	2659100	6482450	V	2002	5, 3
81	Whau Lower	2658691	6479191	V	2005	6, 6
82	Whau Upper	2659738	6476817	M	2004	6, 6
82	Whau Upper	2659738	6476817	M	2005	6, 6
83	Whau Wairau	2658525	6477463	M	2002	4, 1
83	Whau Wairau	2658525	6477463	M	2005	6, 6
84	Whau West	2658343	6479010	M	2005*	6, 6

10 Appendix 2. List of taxa in decreasing order of frequency of occurrence (out of the 95 sample units listed in Appendix 1).

Group	Taxon	Frequency
Polychaete	Polydorid complex	90
Crustacean	*Helice, Hemigrapsus, Macropthalmus*	88
Polychaete	Nereidae	88
Polychaete	*Heteromastus filiformis*	87
Nemertean	Nemertean	87
Crustacean	Phoxocephalidae	87
Mollusc	*Arthritica bifurcata*	79
Polychaete	*Aquilaspio aucklandica*	73
Mollusc	*Nucula hartvigiana*	73
Mollusc	*Austrovenus stutchburyi*	72
Polychaete	*Scolecolepides benhami*	63
Annelida	Capitella, Oligochaetes	62
Crustacean	*Paracalliope novizealandiae*	60
Crustacean	Corophidae	59
Mollusc	*Macomona liliana*	57
Polychaete	*Aricidea sp.*	52
Polychaete	*Glycera* spp.	52
Crustacean	Amphipod other	49
Polychaete	*Cossura consimilis*	49
Crustacean	*Colurostylis* spp.	47
Polychaete	Orbinidae	46
Crustacean	*Halicarcinus* spp.	43
Polychaete	Exogoninae	39
Mollusc	*Theora lubrica*	39
Mollusc	*Cominella glandiformis*	38
Mollusc	*Mactra ovata*	35
Mollusc	*Amphibola crenata*	32
Mollusc	*Notoacmea* spp.	32
Cnidarian	*Anthopleura aureoradiata*	30
Polychaete	*Aonides oxycephala*	30
Polychaete	*Macroclymenella stewartensis*	25
Crustacean	Mysidacea	25
Polychaete	*Pectinaria australis*	24
Crustacean	*Alpheus sp.*	23
Crustacean	*Exosphaeroma* spp.	23
Polychaete	Goniadidae	22
Mollusc	*Paphies australis*	22
Crustacean	Barnacles	21

Group	Taxon	Frequency
Mollusc	*Zeacumantus lutulentis*	17
Polychaete	Cirratulidae	16
Crustacean	Isopod other	16
Polychaete	*Magelona sp.*	16
Crustacean	*Pinnotheres* spp.	15
Mollusc	*Zediloma subrostrata*	15
Polychaete	*Euchone sp.*	14
Polychaete	Paraonidae (not *Aricidea*)	14
Polychaete	*Scolelepis* spp.	13
Crustacean	Tanaidacea	13
Polychaete	*Armandia maculata*	12
Mollusc	Chiton	12
Polychaete	Lepidonotinae	12
Mollusc	*Turbonilla sp.*	12
Mollusc	*Hiatula siliqua*	11
Crustacean	*Waitangi brevirostris*	11
Polychaete	*Aglaophamus macroura*	10
Mollusc	Bivalve unid.	9
Mollusc	*Haminoea zelandiae*	9
Polychaete	Sabellidae	9
Mollusc	*Cominella adspersa*	7
Polychaete	*Paralepidonotus ampulliferus*	7
Polychaete	Syllinae	7
Cnidarian	*Edwardsia sp.*	6
Mollusc	Gastropod unknown	6
Mollusc	*Micrelenchus sp.*	6
Mollusc	*Musculista senhousia*	5
Phoronid	Phoronid	5
Crustacean	*Cyclaspis thomsoni*	4
Crustacean	*Disconatus accolus*	4
Echinoderm	Ophiuroid	4
Mollusc	*Zegaluri tenius*	4
Polychaete	*Asychis amphiglypta*	3
Polychaete	Hesionidae	3
Crustacean	Nebalace	3
Polychaete	Spionidae	3
Echinodermata	*Trochodota dendyi*	3
Mollusc	*Xymene sp.*	3
Mollusc	*Bulla quoyi*	2
Mollusc	Carditidae	2
Crustacean	*Cirolana sp.*	2
Mollusc	*Crassostrea gigas*	2
Mollusc	*Felaniella zelandica*	2

Group	Taxon	Frequency
Polychaete	Lumbrineridae	2
Polychaete	*Owenia fusiformis*	2
Polychaete	*Phyllodocid* spp.	2
Platyhelminth	Platyhelminth	2
Crustacean	*Pontophilus australis*	2
Mollusc	*Solemya parkinson*	2
Mollusc	*Venericardia sp.*	2
Mollusc	*Amalda sp.*	1
Crustacean	Anthuridae	1
Crustacean	*Diastylopsis sp.* (Cumacea)	1
Polychaete	*Harmothoe sp.*	1
Polychaete	Maldanidae	1
Crustacean	Mantis shrimp	1
Polychaete	*Minuspio sp.*	1
Polychaete	*Notomastus sp.*	1
Mollusc	Opistobranch (Philine type)	1
Polychaete	Polynoid	1
Mollusc	*Scintillona zelandica*	1
Sipunculid	Sipunculid	1
Mollusc	*Tellina edgari*	1
Polychaete	*Travisa olens*	1

11 Appendix 3. Brief assessment of sample size bias

The present study used averages of abundance values for the biological samples where replication per site varied from $n = 6$ to $n = 12$, with most sites having $n = 10$ replicates. The primary potential source of bias for multivariate analyses in this case would be caused by underestimates of richness (number of taxa) and therefore overestimates of dissimilarity (e.g., Chao et al. 2005) at sites where fewer replicates were taken. This is because greater replication yields a greater probability of sampling rarer species. Two datasets were generated from available source data: (a) one where averages were recorded for each taxon at each site from $n = 6$ cores and (b) one where averages were recorded for each taxon at each site from all available cores (thus, a mixture of sample sizes). Previous analyses found that the use of averages using all available information (a mixture of sample sizes) resulted in the most useful models (Anderson et al. 2002).

The relationship between the two datasets in terms of richness and in terms of average total abundance was investigated explicitly. There was a slight negative bias for richness using $n = 6$, although not for total average abundances when all samples were used (Fig. A3). The correlations were very strong in each case ($R^2 > 0.90$). Thus, sample size bias was considered negligible, and averages from all available samples at each site were used in order to maximize information content for the development of subsequent models.

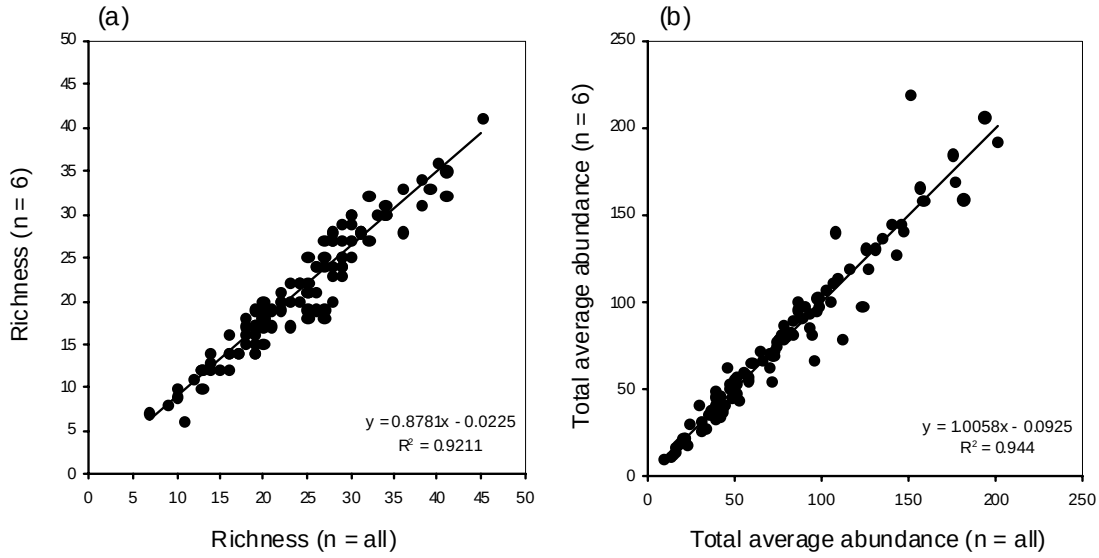
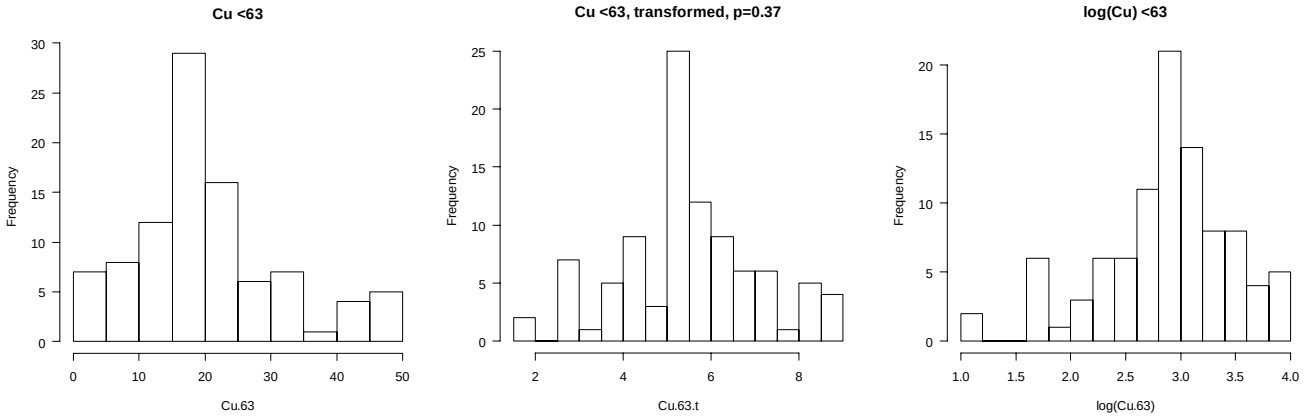


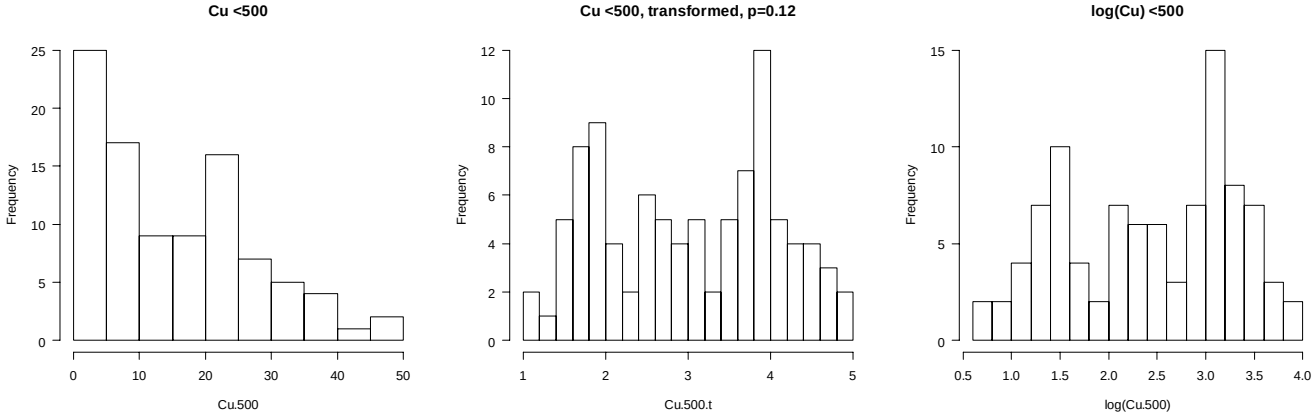
Fig. A3. Relationship between two datasets using n = 6 versus using n = maximum available for (a) richness (total no. of taxa) and for (b) total average abundance per site.

12 Appendix 4. Summary of diagnostics and power transformations trialed for physical and chemical variables.

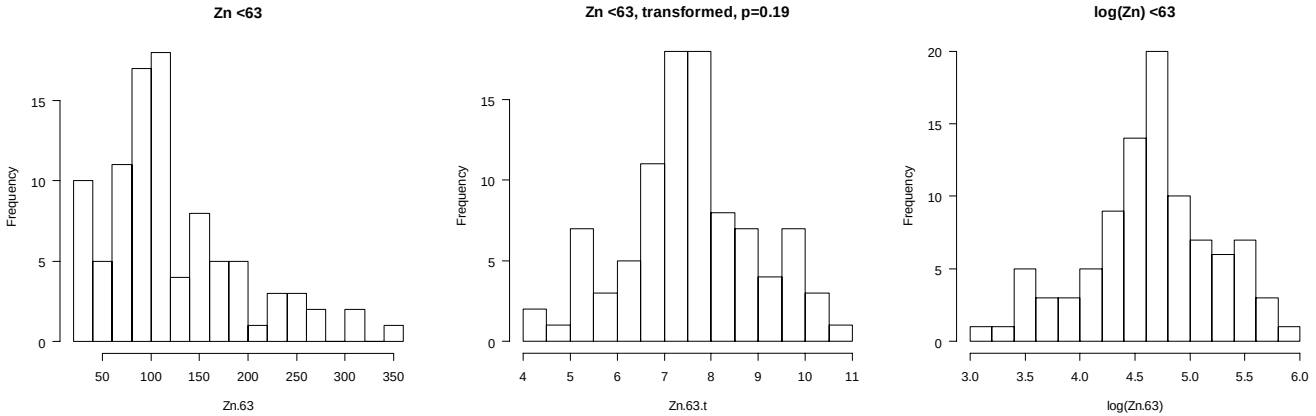
Cu < 63 μm, *power* = 0.37, Shapiro-Wilk test on raw data: *W* = 0.9294, *P* = 7.097e-05, on optimally transformed data: *W* = 0.9776, *P* = 0.1027 and using the log-transformation: *W* = 0.9464, *P* = 0.0006955.



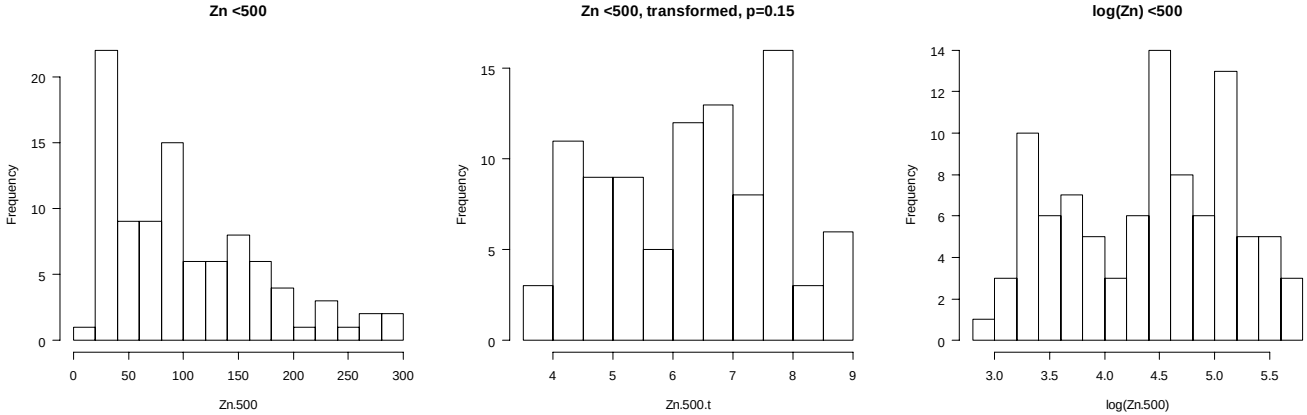
Cu < 500 µm, *power* = 0.12, Shapiro-Wilk test on raw data: *W* = 0.906, *P* = 4.511e-06, on optimally transformed data: *W* = 0.9455, *P* = 0.0006097 and using the log-transformation: *W* = 0.9431, *P* = 0.0004389.



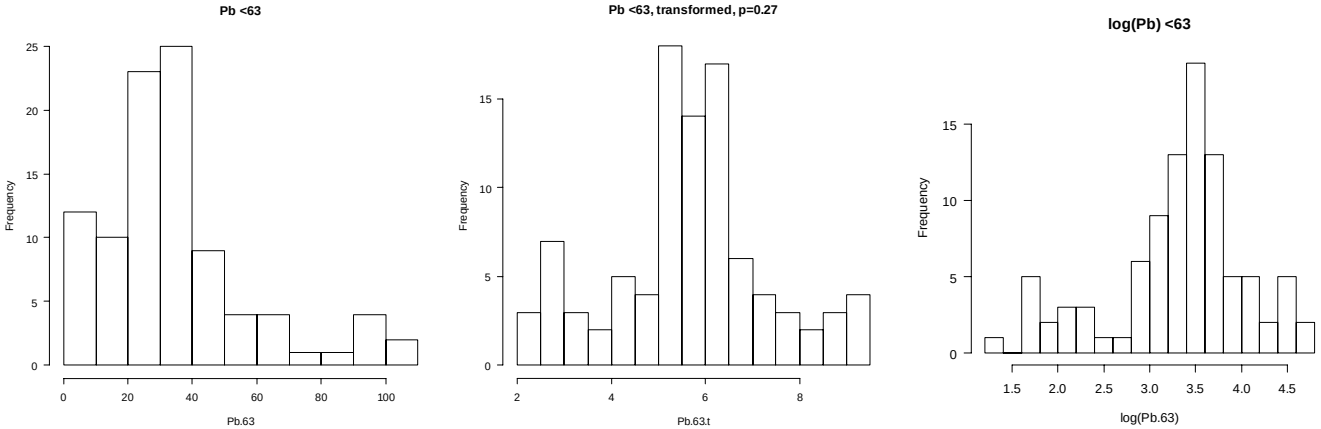
Zn < 63 µm, *power* = 0.19, Shapiro-Wilk test on raw data: *W* = 0.9085, *P* = 5.978e-06, on optimally transformed data: *W* = 0.9842, *P* = 0.3103 and on log-transformed data: *W* = 0.9772, *P* = 0.09524.



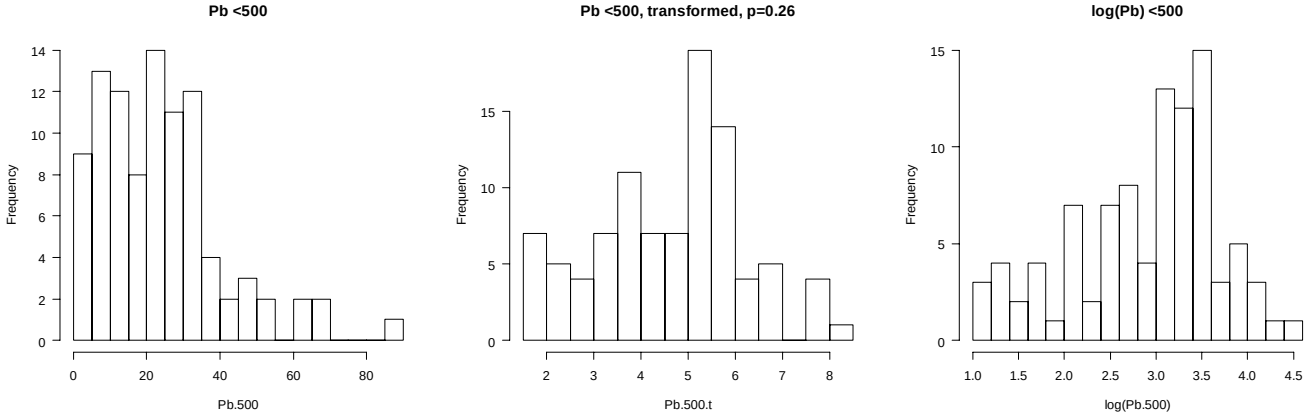
Zn < 500 µm, *power* = 0.15, Shapiro-Wilk test on raw data: *W* = 0.9082, *P* = 5.803e-06, on optimally transformed data: *W* = 0.9587, *P* = 0.004386 and on log-transformed data: *W* = 0.954, *P* = 0.002130.



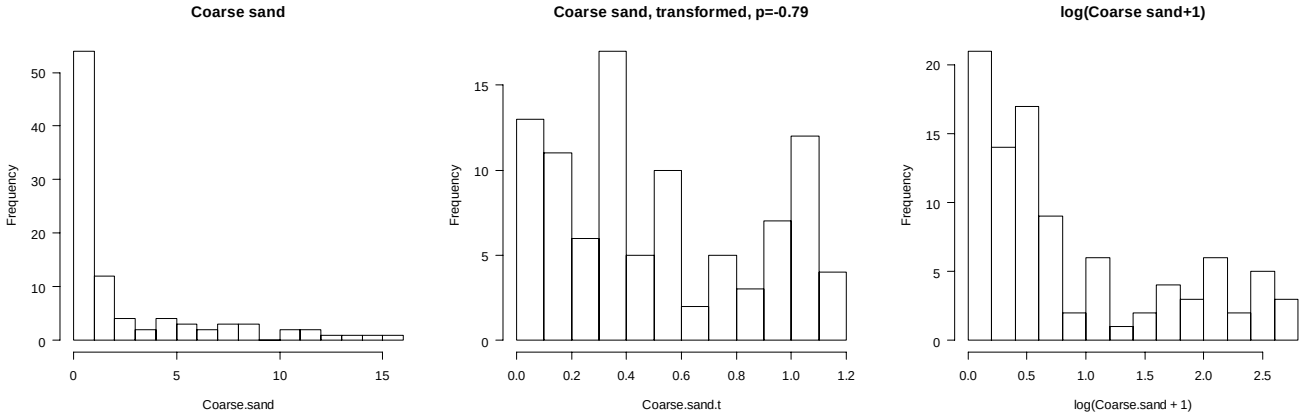
Pb < 63 µm, *power* = 0.27, Shapiro-Wilk test on raw data: *W* = 0.8753, *P* = 2.058e-07, on optimally transformed data: *W* = 0.9674, *P* = 0.01808 and on log-transformed data: *W* = 0.9432, *P* = 0.0004462.



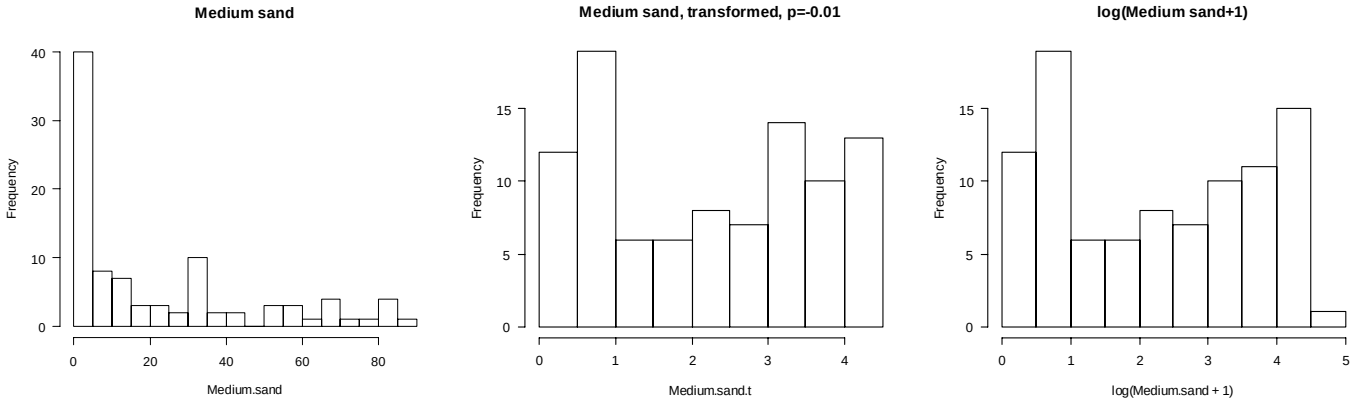
Pb < 500 µm, *power* = 0.26, Shapiro-Wilk test on raw data: *W* = 0.9039, *P* = 3.598e-06, on optimally transformed data: *W* = 0.9763, *P* = 0.08172 and on log-transformed data: *W* = 0.9537, *P* = 0.002053.



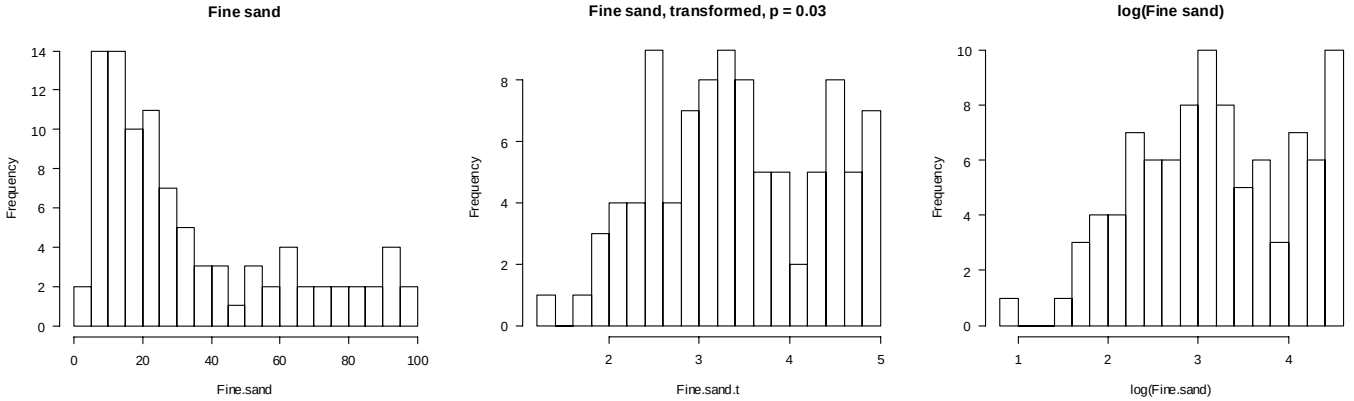
Coarse sand, *power* = –0.79, Shapiro-Wilk test on raw data: *W* = 0.6929, *P* = 8.477e-13, on optimally transformed data: *W* = 0.9192, *P* = 2.024e-05 and on log-transformed data: *W* = 0.8494, *P* = 2.114e-08.



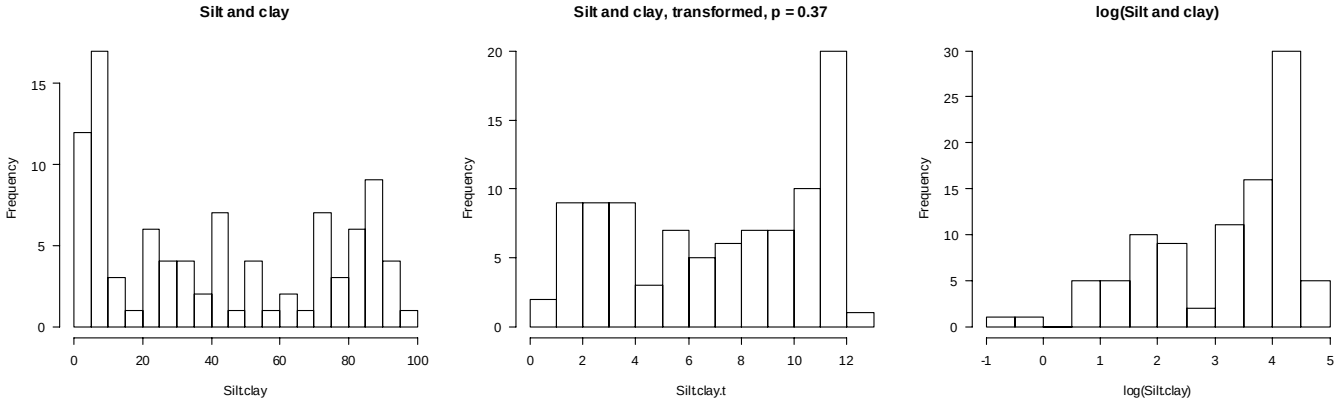
Medium sand, *power* = -0.01, Shapiro-Wilk test on raw data: *W* = 0.7993, *P* = 4.773e-10, on optimally transformed data: *W* = 0.9123, *P* = 9.142e-06 and on log-transformed data: *W* = 0.9123, *P* = 9.154e-06.



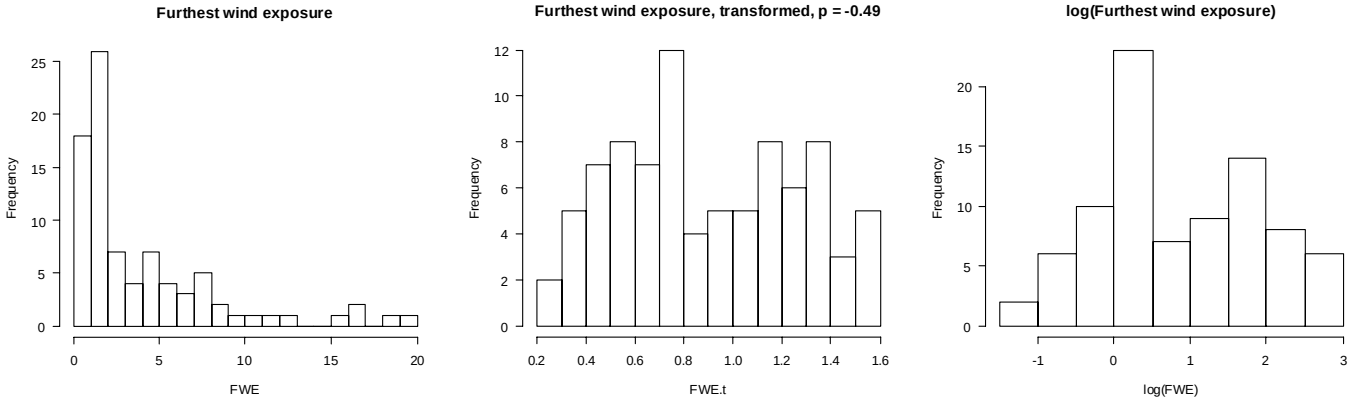
Fine sand, *power* = 0.03, Shapiro-Wilk test on raw data: *W* = 0.8575, *P* = 4.205e-08, on optimally transformed data: *W* = 0.9678, *P* = 0.0194 and on log-transformed data: *W* = 0.9701, *P* = 0.02825.



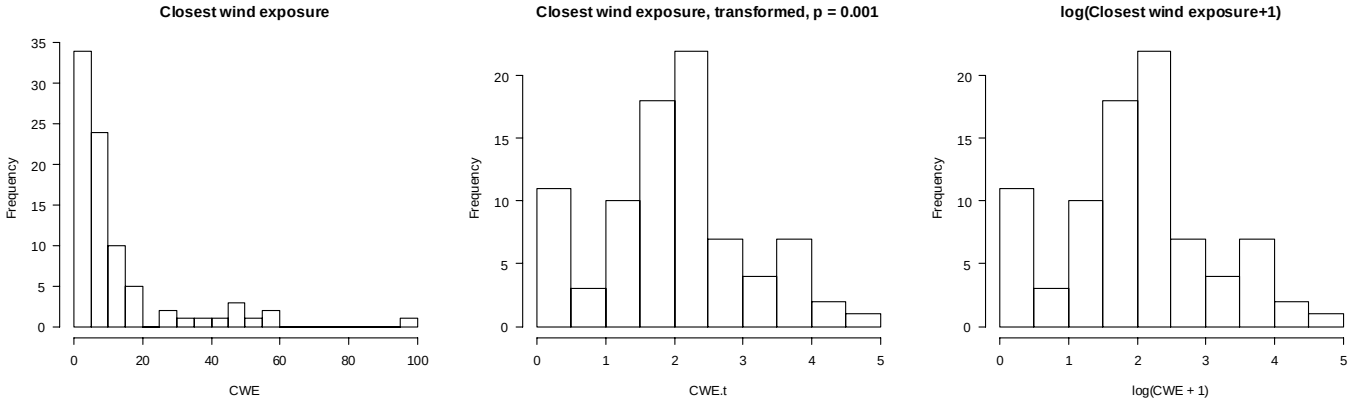
Silt and clay, *power* = 0.37, Shapiro-Wilk test on raw data: *W* = 0.8800, *P* = 3.199e-07, on optimally transformed data: *W* = 0.9055, *P* = 4.275e-06 and on log-transformed data: *W* = 0.8867, *P* = 6.15e-07.



Furthest wind exposure, *power* = -0.49, Shapiro-Wilk test on raw data: *W* = 0.7521, *P* = 1.144e-10, on optimally transformed data: *W* = 0.955, *P* = 0.00481 and on log-transformed data: *W* = 0.9726, *P* = 0.06659.



Closest wind exposure, *power* = 0.001, Shapiro-Wilk test on raw data: *W* = 0.6539, *P* = 7.497e-13, on optimally transformed data: *W* = 0.9606, *P* = 0.01088 and on log-transformed data: *W* = 0.9606, *P* = 0.01082.



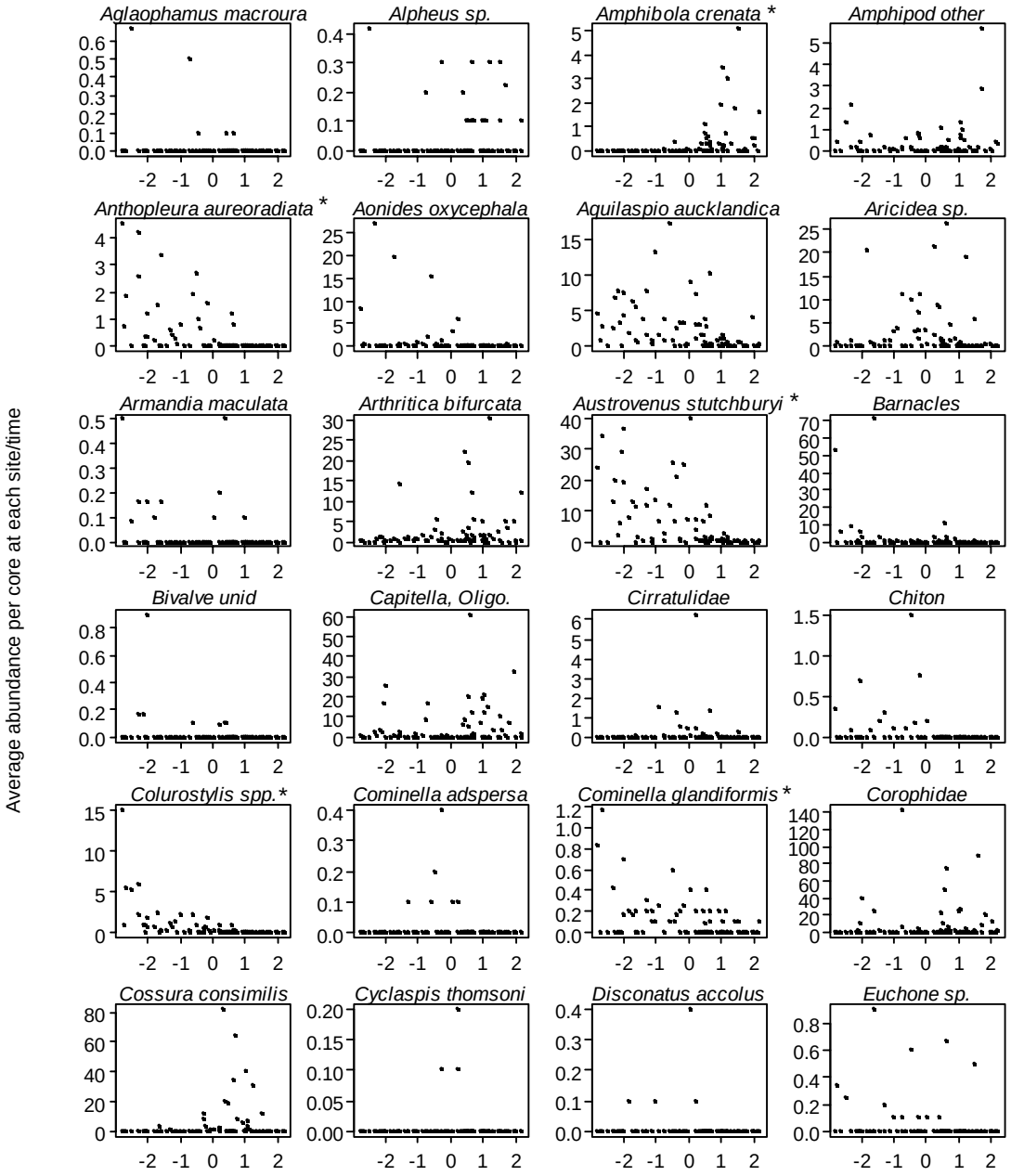
13 Appendix 5. List of taxa in the ecological subset previously proposed and investigated by Anderson et al. (2002) and Hewitt et al. (2005).

Order/Class	Source taxon name	Taxon name as in current list
Amphipoda	Corophidae	Corophidae
Amphipoda	*Methlimedon sp.*	Amphipod other
Amphipoda	Paracalliopidae	*Paracalliope novizealandiae*
Amphipoda	Phoxocephalidae	Phoxocephalidae
Amphipoda	*Torridoharpinia hurleyi*	Phoxocephalidae
Amphipoda	*Waitangi brevirostris*	*Waitangi brevirostris*
Bivalvia	*Arthritica bifurca*	*Arthritica bifurcata*
Bivalvia	*Austrovenus stutchburyi*	*Austrovenus stutchburyi*
Bivalvia	*Hiatula siliqua*	*Hiatula siliqua*
Bivalvia	*Macomona liliana*	*Macomona liliana*
Bivalvia	*Nucula hartvigiana*	*Nucula hartvigiana*
Bivalvia	*Paphies australis*	*Paphies australis*
Bivalvia	*Theora lubrica*	*Theora lubrica*
Crustacea	*Colurostylis lemurum*	*Colurostylis* spp.
Crustacea	*Halicarcinus whitei*	*Halicarcinus* spp.
Crustacea	*Helice crassa*	*Helice, hemigrapsus, macropthalmus*
Crustacea	*Macrophthalmus hirtipes*	*Helice, hemigrapsus, macropthalmus*
Cnidaria	*Anthopleura aureoradiata*	*Anthopleura aureoradiata*
Echinodermata	*Trochodota dendyi*	*Trochodota dendyi*
Gastropoda	*Cominella glandiformis*	*Cominella glandiformis*
Gastropoda	*Notoacmea helmsi*	*Notoacmea spp.*
Isopoda	*Exosphaeroma* spp.	*Exosphaeroma* spp.
Polychaeta	*Aglaophamus macroura*	*Aglaophamus macroura*
Polychaeta	*Aonides oxycephala*	*Aonides oxycephala*
Polychaeta	*Aquilaspio aucklandica*	*Aquilaspio aucklandica*
Polychaeta	*Aricidea sp.*	*Aricidea sp.*
Polychaeta	*Armandia maculata*	*Armandia maculata*
Polychaeta	*Boccardia syrtis*	Polydorid complex
Polychaeta	*Pseudopolydora sp.*	Polydorid complex
Polychaeta	*Heteromastus filiformis*	*Heteromastus filiformis*
Polychaeta	Cirratulidae	Cirratulidae
Polychaeta	*Cossura sp.*	*Cossura consimilis*
Polychaeta	*Euchone sp.*	*Euchone sp.*
Polychaeta	*Glycera sp.*	*Glycera* spp.
Polychaeta	*Goniada emerita*	Goniadidae
Polychaeta	*Lumbrineris* sp.	Lumbrineridae
Polychaeta	*Macroclymenella stewartensis*	*Macroclymenella stewartensis*

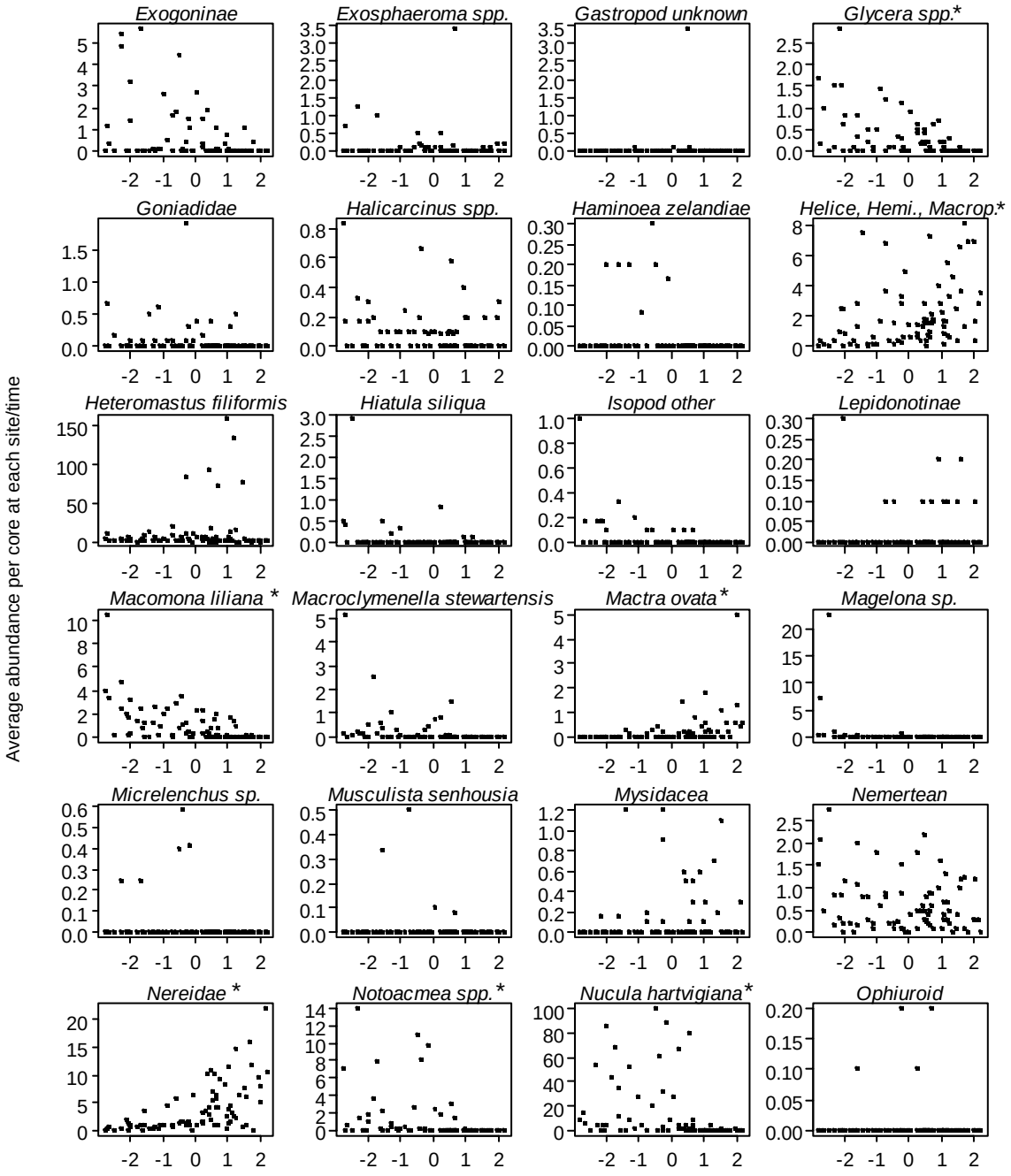
Order/Class	Source taxon name	Taxon name as in current list
Polychaeta	*Magelona ?dakini*	*Magelona sp.*
Polychaeta	Nereidae	Nereidae
Polychaeta	*Orbinia papillosa*	Orbinidae
Polychaeta	*Scoloplos* spp.	Orbinidae
Polychaeta	*Owenia fusiformis*	*Owenia fusiformis*
Polychaeta	Paraonidae	Paraonidae (not *Aricidea*)
Polychaeta	*Pectinaria australis*	*Pectinaria australis*
Polychaeta	*Scolecolepides benhami*	*Scolecolepides benhami*
Polychaeta	*Travisia olens*	*Travisia olens*
Nemertina	Nemerteans	Nemertean
Oligochaeta	Oligochaetes	Capitella, Oligochaetes

14 Appendix 6. Scatterplots of abundances of individual taxa *vs.* PC1.500.

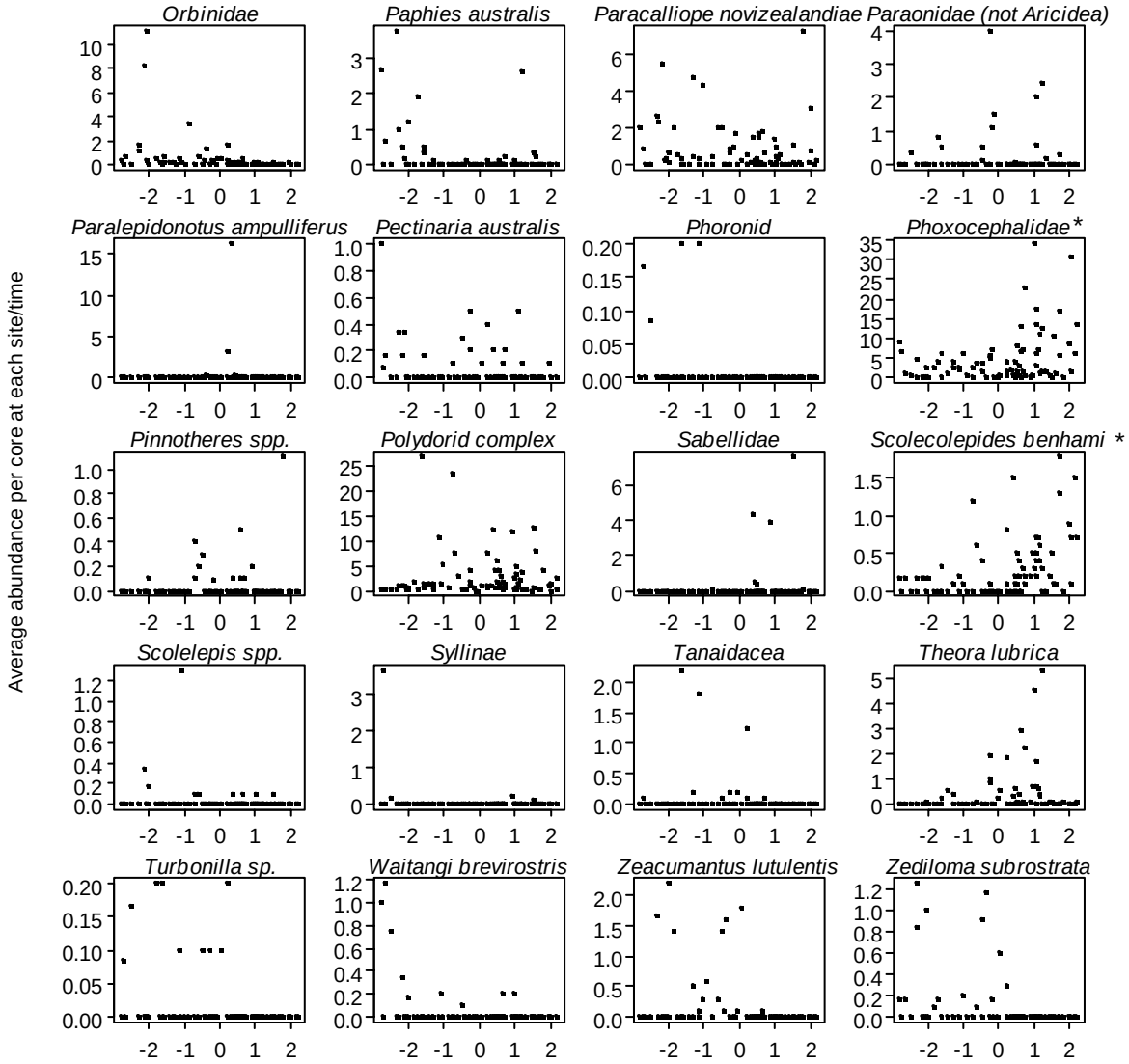
Scatterplots of abundances of individual taxa *vs*. PC1.500. An asterisk (*) denotes a species chosen for subset analysis based on the visual pattern of relationship shown here.



Scatterplots of abundances of individual taxa *vs*. PC1.500. An asterisk (*) denotes a species chosen for subset analysis based on the visual pattern of relationship shown here.

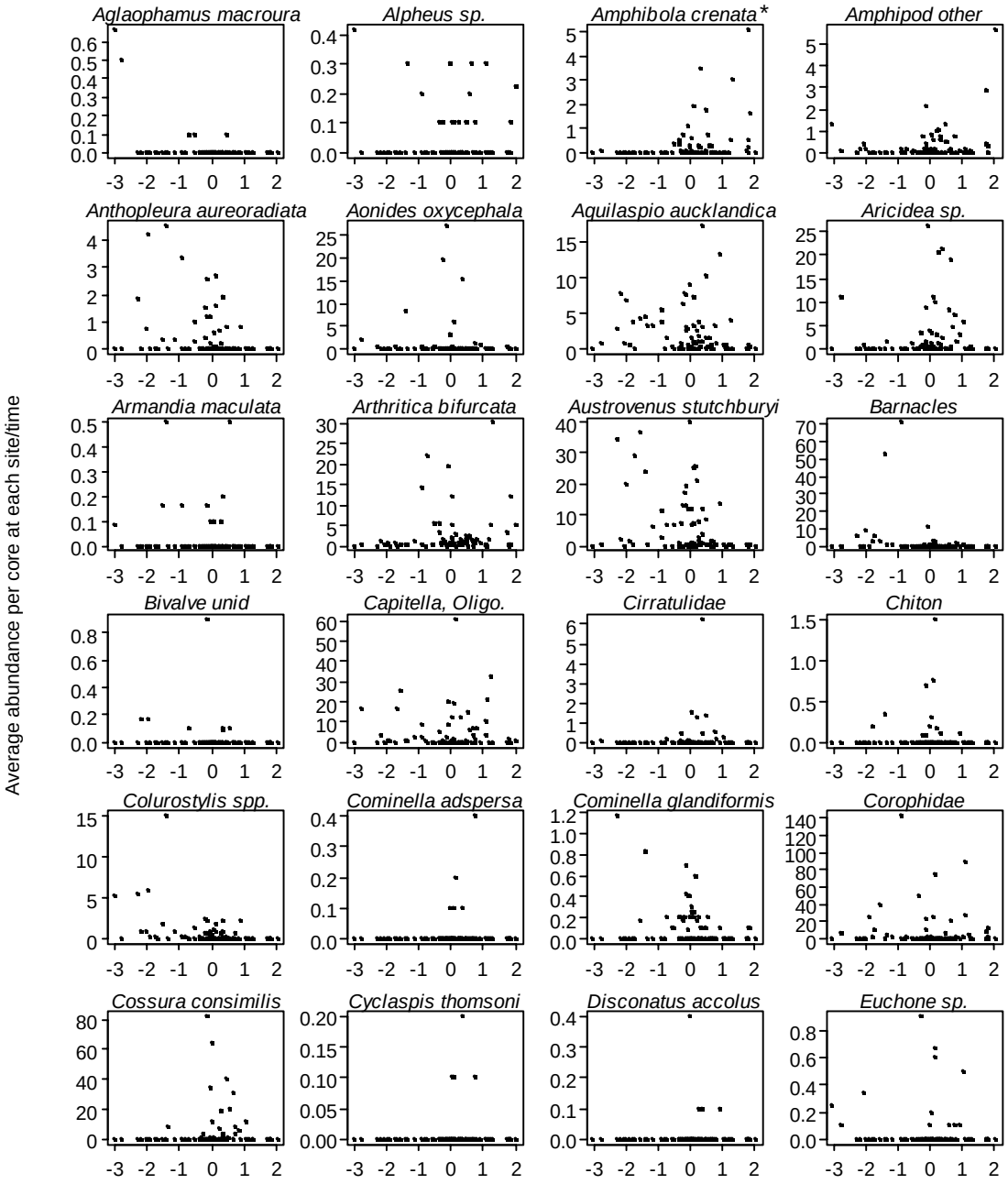


Scatterplots of abundances of individual taxa *vs*. PC1.500. An asterisk (*) denotes a species chosen for subset analysis based on the visual pattern of relationship shown here.

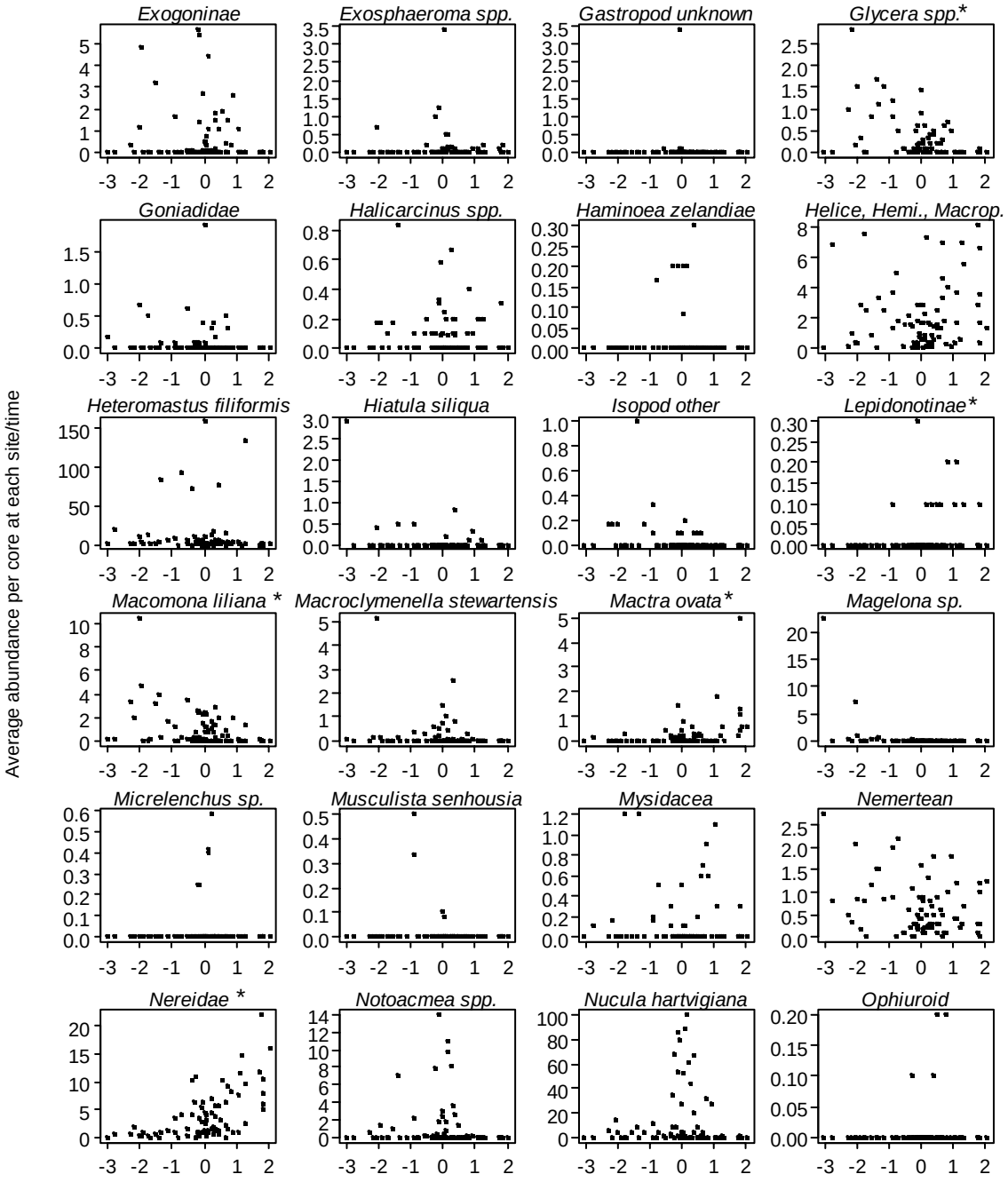


15 Appendix 7. Scatterplots of abundances of individual taxa *vs.* PC1.63.

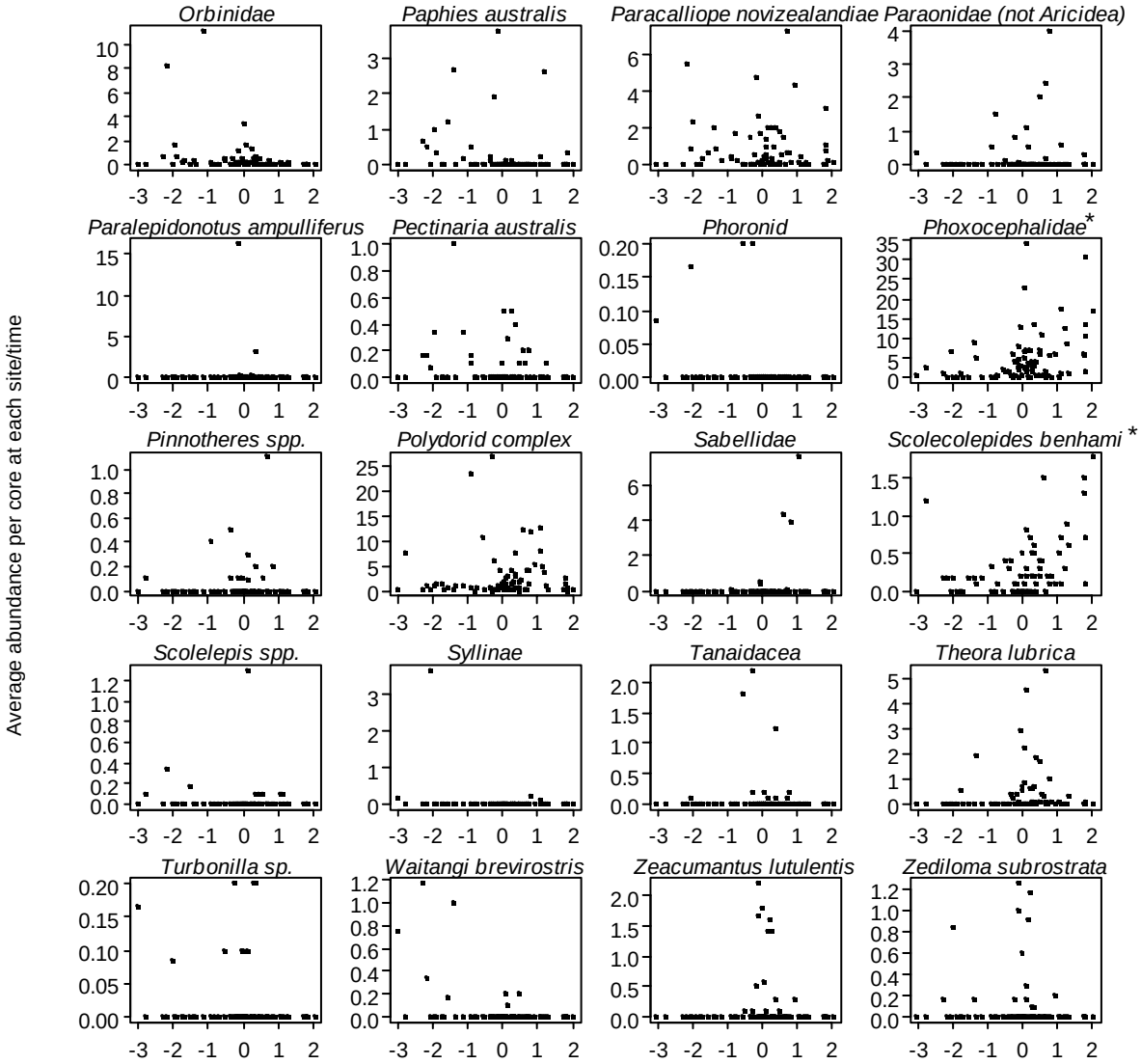
Scatterplots of abundances of individual taxa *vs.* PC1.63. An asterisk (*) denotes a species chosen for subset analysis based on the visual pattern of relationship shown here.



Scatterplots of abundances of individual taxa *vs*. PC1.63. An asterisk (*) denotes a species chosen for subset analysis based on the visual pattern of relationship shown here.



Scatterplots of abundances of individual taxa *vs*. PC1.63. An asterisk (*) denotes a species chosen for subset analysis based on the visual pattern of relationship shown here.



16 Appendix 8. Membership of samples into groups according to physical variables

List of samples indicating membership in one of two groups according to physical variables: F = fine sediments, more sheltered (51 samples) and C = coarser sediments, more exposed (30 samples) and also according to habitats as previously defined (ARC 2002) in terms of either Settling Zone (SZ, 33 samples) or Outer Zone (OZ, 48 samples).

Year	Site.no	Site.name	Mod.Val	Physical Group	Zone
2002	1	Anns Creek	M	F	OZ
2005	1	Anns Creek	M	F	OZ
2002	2	Auckland Airport	V	C	OZ
2004	3	Awatea Rd.	M	F	OZ
2004	4	Bengazai	M	C	OZ
2004	5	Bowden Rd.	M	F	OZ
2005	6	Brigham	M	F	SZ
2002	7	Cape Horn	M	C	OZ
2004	8	Chelsea	M	C	OZ
2002	9	Clarkes Beach	M	C	OZ
2004	10	Coxes, Waitemata	M	C	OZ
2005	11	Coxs	V	C	OZ
2005	12	Glendowie	M	C	OZ
2005	13	Hellyers	M	F	SZ
2005	14	Hellyers outer	M	F	OZ
2002	15	Henderson Entrance	M	C	OZ
2004	15	Henderson Entrance	M	C	OZ
2004	16	Henderson lower	M	F	SZ
2002	17	Henderson Upper	M	F	SZ
2005	17	Henderson Upper	V	F	SZ
2005	18	Herald Island	V	F	OZ
2005	19	Hi North	M	C	OZ
2004	20	Hillsborough	M	C	OZ
2005	21	Hobson: Purewa Bridge	V	F	SZ
2005	22	Hobson: Tohunga	M	F	OZ
2002	23	Hobsonville	M	C	OZ
2005	23	Hobsonville	M	C	OZ
2005	24	Kaipatiki	M	F	SZ
2004	25	Kendalls	V	C	OZ
2005	26	Little Shoal Bay	V	C	OZ
2005	27	Lower Shoal Bay	M	C	OZ
2005	28	Lucus outer	M	C	OZ
2004	29	Lucus Te Wharau	M	F	SZ
2005	30	Lucus Upper	M	F	SZ
2004	31	Mangemangeroa B	M	C	OZ
2004	32	Mangemangeroa E	V	C	OZ
2005	33	Mangere Cemetery	M	F	OZ
2005	34	Mangere Inlet: Harania Creek	M	F	OZ
2005	35	Mangere Inlet: Kiwi Esplanade	M	F	OZ
2005	36	Mangere Inlet: Tararata Creek	V	F	OZ
2002	37	Meola Inner	M	F	SZ

Year	Site.no	Site.name	Mod.Val	Physical Group	Zone
2005	37	Meola Inner	M	F	SZ
2004	38	Meola Outer	M	C	OZ
2002	39	Meola Reef	M	C	OZ
2005	39	Meola Reef	M	C	OZ
2005	40	Meola West	M	F	OZ
2002	41	Middlemore	M	F	SZ
2005	41	Middlemore	M	F	SZ
2002	42	Motions	M	F	SZ
2005	42	Motions	M	F	SZ
2005	43	Motions East	M	C	OZ
2005	44	Newmarket	V	C	OZ
2005	45	Ngataringa Bay	M	F	SZ
2005	46	Oakley	M	F	SZ
2004	47	Okura D	M	C	OZ
2004	48	Okura J	V	C	SZ
2004	49	Orewa F	M	F	OZ
2004	50	Orewa G	M	F	OZ
2004	51	Otahuhu creek	M	F	SZ
2005	52	Out Main UWH	M	C	OZ
2005	53	Pakuranga	V	F	SZ
2005	54	Pakuranga mid	M	F	SZ
2004	55	Panmure	M	F	SZ
2005	56	Paremoremo	M	F	SZ
2005	57	Paremoremo upper	M	F	SZ
2005	58	Pollen Island	M	C	OZ
2004	59	Princess St	M	C	OZ
2005	60	Puhinui	M	F	SZ
2002	61	Puhinui, Entrance	M	C	OZ
2004	62	Puhoi F	M	C	OZ
2004	63	Puhoi H	M	F	OZ
2005	64	Pukaki	M	F	SZ
2004	65	Purewa	M	F	SZ
2005	66	Rangitopuni	M	F	SZ
2005	67	Rangitopuni UWH	M	F	SZ
2004	68	Shoal Bay, Hillcrest	M	F	SZ
2004	69	Shoal Bay, Upper	M	C	SZ
2004	70	Turanga G	M	C	OZ
2004	71	Turanga J	M	F	SZ
2005	72	Upper main UWH	M	F	OZ
2004	73	Victoria Ave	M	F	OZ
2004	74	Waiwera E	M	F	OZ
2004	75	Waiwera J	M	F	OZ
2005	76	Weiti	M	F	SZ
2002	77	Whakataka	M	C	OZ
2005	77	Whakataka	M	C	OZ
2005	78	Whau East	M	F	OZ
2004	79	Whau Entrance	M	C	OZ
2002	80	Whau Entrance WHO A	V	C	OZ
2005	81	Whau Lower	V	F	OZ
2004	82	Whau Upper	M	F	SZ
2005	82	Whau Upper	M	F	SZ
2002	83	Whau Wairau	M	F	SZ

Year	Site.no	Site.name	Mod.Val	Physical Group	Zone
2005	83	Whau Wairau	M	F	SZ
2005	84	Whau West	M	F	OZ

17 Appendix 9. Summary of CAP analyses relating biotic assemblages to pollution gradients

Summary of CAP analyses relating biotic assemblages to pollution gradients based on either the whole sample (< 500 μm) or on the mud fraction (< 63 μm), done separately for samples from either the Settling Zone (SZ, 33 samples) or the Outer Zone (OZ, 48 samples). Table headings are as given for Tables 13 and 17 in the text.

<500 μm

	Set	*m*	prop.G	SS_{RES}	δ_1	correl
BC, sqrt	SZ	5	0.686	0.653	0.557	0.746
Euc, ln(x+1)	SZ	10	0.904	0.598	0.738	0.859
Mod. Gower	SZ	7	0.683	0.634	0.546	0.739
BC, sqrt	OZ	9	0.825	0.400	0.711	0.843
Euc, ln(x+1)	OZ	11	0.864	0.545	0.668	0.817
Mod. Gower	OZ	17	0.875	0.466	0.775	0.880

< 63 μm

	Set	*m*	prop.G	SS_{RES}	δ_1	correl
BC, sqrt	SZ	6	0.745	0.688	0.466	0.683
Euc, ln(x+1)	SZ	11	0.924	0.477	0.790	0.889
Mod. Gower	SZ	8	0.727	0.668	0.519	0.720
BC, sqrt	OZ	15	0.943	0.355	0.815	0.903
Euc, ln(x+1)	OZ	11	0.864	0.446	0.741	0.861
Mod. Gower	OZ	28	0.983	0.400	0.896	0.947