Technical Report 432
SFEI Contribution 89
April 2004

Evaluation of Benthic Assessment Methodology in Southern California Bays and San Francisco Bay

Ananda Ranasinghe
Bruce Thompson
Robert Smith
Sarah Lowe
Kenneth Schiff

Southern California Coastal Water Research Project

Evaluation of Benthic Assessment Methodology in Southern California Bays and San Francisco Bay

J. Ananda Ranasinghe[1]
Bruce Thompson[2]
Robert W. Smith[3]
Sarah Lowe[2]
Kenneth C. Schiff[1]

[1]Southern California Coastal Water Research Project
7171 Fenwick Lane
Westminster, CA 92683
www.sccwrp.org

[2]San Francisco Estuary Institute
7770 Pardee Lane
Oakland, CA 94621

[3]P.O. Box 1537
Ojai, CA 93024

April 2004

Southern California Coastal Water Research Project
Technical Report 432

San Francisco Estuary Institute
Contribution 89

ACKNOWLEDGEMENTS

The California State Water Resources Control Board supported this study through Agreement Number 01-190-250-0. We are grateful to our Project Manager for the State Water Resources Control Board, Dominic Gregorio, and Steve Bay and Steve Weisberg of SCCWRP for their direction, guidance, and review. Chris Beegan recognized the need to evaluate benthic condition on the same scale throughout California and initiated this project to support that mission.

Thanks also to Shelly Moore, map preparation; Valerie Raco-Rands, cover composition; Linda Fuller, editing; and Debbie Elmore, desktop publishing.

The data used in this report were produced in southern California and San Francisco Bay for many projects funded by several agencies. The samples contributing data were also collected by many field crews and processed by many sorters and taxonomists. While space limitations do not allow us to acknowledge every participant by name, the authors thank each and every one.

ABSTRACT

Benthic invertebrates are important components of aquatic food webs that are sensitive to changes in environmental condition, which makes them ideal candidates for monitoring. However, assessments of benthic community data are difficult because biological communities are complex. Dozens of species and hundreds of organisms are often found in a single sample, with numbers and species of organisms varying from habitat to habitat. To improve environmental assessments, scientists have developed biological indices that reduce complex data to single values useful for evaluating community health using thresholds of concern and for tracking trends in benthic condition. In California, biological indices with distinctly different approaches have been designed for use in two specific regions. A Benthic Response Index (BRI) was developed recently for the bays and harbors of southern California and an Index of Biological Integrity (IBI) was developed recently for San Francisco Bay. This study develops a BRI in San Francisco Bay and an IBI in southern California and compares results for the two benthic assessment methods at the same sampling sites.

A BRI was developed for San Francisco Bay using 165 samples from two different habitats. Pollution tolerance scores were calculated for species occurring in two or more samples in each habitat (59 species in the northern bay and 183 species in the southern bay). Five categories of impact were recognized based on a gradient of species loss; thresholds were established where 5, 25, 50, and 80% of the species present at the peak were lost.

An IBI was developed in southern California using 155 samples from two habitats. Twenty-two indicators were evaluated for IBI development. Four were selected for the northern habitat (including total number of individuals, total number of taxa, total number of *Dorvillea (Schistomeringos)* sp., and total number of *Spiophanes duplex*); and three for the southern habitat (including total number of individuals, total number of taxa, and total number of molluscan taxa). After identifying ranges of values observed for these indicators at reference sites in each habitat, assessment values of 1 (if outside the range) or 0 (if within the range) were assigned to each indicator at each site. If a site had a cumulative assessment value of 2 or more, that site was considered to be impacted.

Comparisons between the BRI and IBI assessments indicated that, despite the differences in development approach, both indices produced similar results. There was 78.1% agreement in site classifications for southern California and 67.6% agreement in site classifications for San Francisco Bay. Where there were differences, the BRI categorized more sites as disturbed in both regions when the IBI did not.

The new assessment tools in San Francisco Bay and southern California need additional refinement and validation. Both methods are limited by three factors: (1) lack of independent data for validation; (2) insufficient data from highly disturbed sites to define the entire range of the impact gradient; and (3) uncertainty in the effect of environmental variables that can affect assemblage composition regardless of pollution impacts.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Benthic index-based approaches to summarizing data (Engle *et al.* 1994, Weisberg *et al.* 1997, Engle and Summers 1999, Van Dolah *et al.* 1999, Paul *et al.* 2001, Smith *et al.* 2001, Llanso *et al.* 2002, Smith *et al.* 2003) have facilitated the use of benthic infauna as indicators of environmental condition in marine and estuarine environments (Hyland *et al.* 1999, Bergen *et al.* 2000, Dauer *et al.* 2000, Summers 2001, Hyland *et al.* 2003). While reducing complex biological data to a single value has disadvantages, the resulting indices remove much of the subjectivity associated with interpreting data. The indices also provide a simple means of communicating complex information to managers and correlating benthic responses with stressor data (Dauer *et al.* 2000, Hale *et al.* 2004).

In California, benthic indices have been successfully developed to assess bay and estuarine habitat condition in two regions using different approaches. Smith *et al.* (2003) developed a Benthic Response Index (BRI) for southern California bays while Thompson and Lowe (In press) developed the San Francisco Bay Index of Biotic Integrity (IBI). The BRI is a multivariate approach originally developed on the mainland shelf of southern California (Smith *et al.* 2001). It uses an abundance-weighted pollution tolerance score to distinguish multiple levels of impact ranging from reference to loss in ecosystem function. The pollution tolerance score was developed based on the response of individual species to pollution gradients; most species in a sample are included when calculating a BRI value for a site. The IBI, well established in freshwater (Karr and Chu 1999), uses a set of benthic indicators (e.g., species diversity, abundance of key taxa) in a multi-metric index to distinguish impacted from reference benthic conditions. Adaptations of the IBI have also been applied in marine and estuarine areas along the eastern coast of the United States (Weisberg *et al.* 1997, Van Dolah *et al.* 1999, Llanso *et al.* 2002) and in California's estuaries for the Bay Protection and Toxic Clean-up Program (BPTCP: Jacobi *et al.* 1998, Anderson *et al.* 2001, Hunt *et al.* 2001). The San Francisco Estuary IBI method used in this study and the IBI method used in the BPTCP differ in the way reference conditions were established and in the choice of benthic indicators.

Although these two benthic indices are useful for assessments within their respective regions, assessment results cannot be compared between regions. At present, the BRI and IBI are of limited use to the State of California's Water Resources Control Board (SWRCB), which is responsible for protecting resources throughout California. Before living aquatic resources in California bays and harbors can receive equal levels of protection, a method to evaluate benthic condition in both regions on the same scale must be developed.

The objective of this study was to compare the scoring of benthic condition used by the BRI and IBI assessment methods. A BRI was developed in San Francisco Bay and BRI results were compared with the original IBI results. Likewise, an IBI was developed in southern California and IBI results were compared with the original BRI results. This report is divided into three sections: Section 1 presents the methods, results, and discussion for San Francisco Bay. Section 2 presents the methods, results, and discussion for southern California. Section 3 summarizes our conclusions.

DEVELOPMENT OF A BENTHIC RESPONSE INDEX FOR SAN FRANCISCO BAY AND COMPARISON WITH SAN FRANCISCO BAY IBI ASSESSMENTS

Methods

(i) Development of a Benthic Response Index for San Francisco Bay

The Benthic Response Index (BRI) is the abundance-weighted average pollution tolerance of species occurring in a sample (Smith *et al.* 2001, 2003). The general index formula is:

$$BRI_s = \frac{\sum_{i=1}^{n} a_{si}^{f} p_i}{\sum_{i=1}^{n} a_{si}^{f}} \quad (1)$$

where BRI_s is the BRI value for sampling unit s, n is the number of species in s, p_i is the pollution tolerance of species i, a_{si} is the abundance of species i in s, and f is an exponent used to transform the abundance values.

The primary objective of BRI development is to assign pollution tolerance scores p_i to species based on their position on a pollution gradient. Once assigned, the scores can be used to assess the condition of the benthic community by calculating the BRI. A six-step process was used to assign and validate pollution tolerance scores for benthic infauna in San Francisco Bay.

1. Data were assembled from 10 projects distributed throughout San Francisco Bay (Figure 1, Tables 1 and 2) and adjusted for compatibility.

2. San Francisco Bay was divided into two habitats, the northern bay and the southern bay, based on differences in naturally occurring benthic assemblages identified by cluster analysis. The northern bay and southern bay were divided at 37°54' (approximately the southern tip of Brooks Island, just south of Richmond Inner Harbor). The index was developed separately in each habitat because the numbers and kinds of benthic organisms vary naturally, and comparisons to determine altered states should vary accordingly.

3. An ordination analysis was performed to quantify species changes along environmental gradients, and a pollution vector was identified to quantify species changes along the pollution gradient. In ordination analysis, samples are displayed as points in a multi-dimensional space where the distance between points is proportional to differences in species composition among the samples. Different environmental gradients causing species change often correlate with vectors extending in different directions in this space. The pollution vector was defined as the direction maximally correlated with two indicators of potential pollution effects: (1) the mean effects range median (ERM) quotient, which is an integrated measure of chemical contamination in the sediments, and (2) the acute toxicity of the sediments to amphipods.

4. For each species, a pollution tolerance score was calculated as the weighted-average position of its abundance distribution along the pollution vector. Pollution tolerance scores were calculated for each species occurring in two or more samples in each habitat. The pollution vectors were normalized to a scale of 0-100 that was equivalent among habitats.

5. To give index values an ecological context, four thresholds of biological response to pollution were identified. A reference threshold was identified below which natural benthic assemblages normally occur, and three thresholds of response to disturbance were identified that were equal to the thresholds developed for the southern California coast (Smith *et al.* 2001) and bays (Smith *et al.* 2003) BRI.

6. Finally, the index was validated internally by comparing index values at each site with data indicating potential pollution effects. In addition, relationships between the index values and several habitat variables were examined to ensure that the index was measuring the pollution gradient as intended, rather than habitat gradients.

The details for each step are provided below.

1. Assemble Data

Benthic species abundance data for 1,153 samples from 10 sources that sampled in the San Francisco Bay estuary were combined to create the project database (Table 1). Whenever available, data about chemical contaminant concentrations, toxicity of the sediments to amphipods, and habitat measures such as bottom depth, sediment grain size composition, and total organic carbon were included, provided they were collected at the benthic sampling sites at the same time (Table 2). These data were collected throughout the year; they were not limited to summer samples as in southern California (Smith *et al.* 2001, 2003). They were gathered over several years (Table 1) from many regions of San Francisco Bay (Figure 1).

The data sources all identified benthic organisms retained in sieves to the lowest practical taxon, most often species, and counted them. Taxonomic inconsistencies among programs were eliminated by cross-correlating species lists to identify differences in nomenclature or taxonomic level, consulting taxonomists from each program, and resolving discrepancies. In a few cases, multiple taxa were combined to resolve taxonomic inconsistencies in the data. Most samples were collected with 0.05 m^2 benthic grabs; abundances in the others were normalized to an area of 0.05m^2. Mean abundances were used when multiple 0.05m^2 pseudoreplicates were collected at a site.

Mean ERM quotients (Long and MacDonald 1998) were calculated as an integrated measure of chemical contamination at each site if data about concentrations of contaminants were available. The ERM value for each contaminant is the level at which biological effects are likely (Long *et al.* 1995), and the mean ERM quotient is the mean ratio of observed concentrations to the ERM values. Depending on available data, ERM quotients were calculated for 16 to 24 of the contaminants for which ERM values exist (Thompson *et al.* 2000). The contaminants included 8 trace metals (arsenic, cadmium, chromium, copper, lead, mercury, silver, and zinc); 13 polycyclic aromatic hydrocarbons or PAHs (acenaphthene, acenaphthylene, anthracene, benzo[a]pyrene, benzo[a]anthracene, chrysene, dibenz[a,h]anthracene, fluoranthene, fluorine, 2-methylnaphthalene, naphthalene, phenanthrene, and pyrene); p'p'-DDE; total DDTs; and total polychlorinated biphenyls (PCBs). The contaminant concentrations were measured using comparable laboratory analysis methods for all the samples.

Amphipod toxicity test results were expressed as the mean control-adjusted amphipod mortality for each site. Sediment toxicity to amphipods (*Eohaustorius estuarius* or *Ampelisca abdita*) was measured by a 10-day acute toxicity test (Heitmuller *et al.* 1999, Thompson *et al.* 1999, Hunt *et al.* 2001).

2. Identify Habitats with Distinct Natural Assemblages

A process similar to Bergen *et al.* (2001) was used to identify naturally occurring benthic assemblages in San Francisco Bay and the habitat factors that structure them. After eliminating potentially contaminated sites, assemblages were identified using hierarchical cluster analysis and tested habitat variables across dendrogram splits to assess whether the assemblages occupied different habitats.

Because the objective was to define natural groupings of samples with similar species composition, screening criteria similar to those of Bergen *et al.* (2001) were used to eliminate potentially contaminated sites from the analysis. A sample was considered potentially contaminated if the mean ERM quotient was more than 0.1 (Long and MacDonald 1998). Only benthic data from samples with synoptic chemical contaminant data were selected.

The benthic data were also restricted to samples that were collected using gear sampling areas of, or close to, 0.05 m^2 and screened through 0.5mm sieves. Abundance data from samples collected with an 0.044m^2 Young Grab (Table 1) were normalized to 0.05m^2; only four of these samples survived the elimination process described above. After screening, 86 samples (Figure 2) from 23 sites (Figure 3) were available for cluster analysis.

Cluster analysis was conducted using flexible sorting of Bray-Curtis dissimilarity values with β=-0.25 (Bray and Curtis 1957, Lance and Williams 1967, Clifford and Stephenson 1975). For station (Q-mode) analyses, abundances were square-root transformed and then standardized by the species mean of values higher than zero to reduce the influence of dominant species (Smith 1976, Smith *et al.* 1988). The step-across distance re-estimation procedure (Williamson 1978, Bradfield and Kenkel 1987) was applied to dissimilarity values over 0.80 to reduce the distortion of ecological distances caused by joint absences of a high proportion of species. The distortion occurs due to the common non-monotonic truncated nature of species distributions along environmental gradients (Beals 1973). For species (R-mode) analysis, the square-root transformed abundance data were standardized by the species minimum. Prior to cluster analysis, species contributing little information were excluded by eliminating species occurring in fewer than 5 samples unless the total abundance in all 86 samples was more than 50 individuals (Smith 1976). Of 217 taxa in the original data, 159 taxa were included in the analysis.

The number of habitat-defined assemblages was determined by sequentially examining each split of the cluster analysis dendrogram, starting at the top, to assess whether each split reflected habitat differentiation. Habitat differentiation was defined as a significant (Mann-Whitney U-test) difference in habitat variables between the sets of samples defined by the dendrogram split and segregation of more than 90% of the samples in the split by the significant habitat variables. Six continuous variables were tested (salinity, depth, fine sediment content, total organic carbon, latitude, and longitude). This process was conducted along each branch of the dendrogram until a split yielded no significant difference or a split contained fewer than 10 samples. Probabilities were not adjusted to account for multiple testing because we were only interested in controlling the comparison-wise error rate.

3. Identify the Pollution Vector in Ordination Space

A total of 167 samples processed with 0.5 mm sieves and with synoptic sediment contaminant concentration and amphipod toxicity data were selected for index development (Table 1). The data selected were collected throughout the year and were not limited to summer samples as for the southern California BRIs (Smith *et al.* 2001, 2003). Samples were gathered over several years (Table 1) and from many regions of the San Francisco Bay estuary.

Using these data, gradients of species change caused by environmental gradients were quantified using principal coordinates ordination analysis (Gower 1966, Pielou 1984) on a Bray-Curtis dissimilarity matrix (Bray and Curtis 1957). A single analysis was used for all the selected data. Before calculating the dissimilarity matrix, abundances were square root transformed and standardized by the species mean of values higher than zero, to reduce the influence of dominant species (Smith 1976, Smith *et al.* 1988). The step-across distance re-estimation procedure (Williamson 1978, Bradfield and Kenkel 1987) was applied to dissimilarity values over 0.80 to reduce the distortion of ecological distances caused by joint absences of a high proportion of species. The distortion occurs due to the common non-monotonic truncated nature of species distributions along environmental gradients (Beals 1973). All species occurring at two or more sites were included for calculation of the dissimilarity matrix.

Next, canonical correlation analysis (Cooley and Lohnes 1971, Gittins 1979, Dillon and Goldstein 1984) was used to find directions (gradients or vectors) of species change in the ordination space that maximally correlated with two pollution indicator variables. Specifically, the canonical correlation compared the first 20 ordination axes with the mean ERM quotient (Long and MacDonald 1998) and the mean control-adjusted mortality in the amphipod toxicity tests. The overall pollution vector was calculated as the average direction between the ERM quotient and amphipod toxicity vectors. Separate canonical correlation analyses were performed for the northern and the southern bay data. A simple example of our method is included in Appendix A.

4. Identify the Position of Each Species on the Pollution Vector

The pollution tolerance score for each species was defined as its abundance-weighted average position on the pollution vector. For each species p_i, the pollution tolerance score was calculated as:

$$p_i = \frac{\sum_{j=1}^{t} a_{ij}^{e} g_j}{\sum_{j=1}^{t} a_{ij}^{e}} \quad (2)$$

where e is an exponent for transforming the abundance, and t is the number of samples to be used in the sum, with only the samples with highest t species abundance values included. The g_j is the position of the j^{th} sample on the pollution gradient.

An optimization procedure was used to find values for the unspecified parameters f, e, and t in Equations 1 and 2. The optimization consisted of computing correlation coefficients (r_{I_s,g_s}) between BRI index values for each sample and the position of the sample on the pollution vector for all combinations of e = 0, 1, 0.5, 0.33, 0.25, and t = 1 to 100 in Equation 2, and f = 0, 1, 0.5, 0.33, 0.25 in Equation 1. The combination of f, e, and t values that maximized the correlation coefficient was chosen and substituted in the general formulae to calculate pollution tolerance scores for each species and BRI index scores for each sample. The procedure was applied separately to the data in each habitat defined in step two. To avoid the higher sampling error associated with rare species, p_i values were computed only for species occurring two or more times in a data subset.

To enhance interpretability, index values for all the data were standardized to a 0 to 100 scale by (1) translating the index scale for each habitat so that the minimum index value was equal to zero, and (2) rescaling the translated values so that index values ranged from 0 to 100 in both habitats. To preserve the relative scale of differences along the pollution gradient in the ordination space, rescaling was performed after combining all the data from both habitats.

5. Develop Assessment Thresholds

To give index values an ecological context and facilitate their interpretation and use for evaluation of benthic community condition, a reference threshold and three thresholds of response to disturbance were defined. These thresholds were equivalent to the thresholds established for the southern California mainland shelf and bay BRIs (Smith *et al.* 2001, 2003). The goal was to define the reference threshold as a value toward the upper end of the range of index values for sites that had minimal known anthropogenic influence. It was established at the point on the pollution vector where pollution effects first resulted in a net loss of species.

The other three thresholds involved defining increasing levels of deviation from the reference condition. These thresholds were based on determinations of index values at which 25%, 50%, and 80% of the species present at the reference threshold were lost. These intervals are referred to as Response Levels 1, 2, 3, and 4 and indicate increasing levels of disturbance (Smith *et al.* 2001, 2003). Response Level 1 indicates marginal disturbance while Response Levels 2, 3, and 4 indicate clearly disturbed benthic communities.

6. Validate Index Values and Pollution Tolerance Scores

The index was validated internally by comparing index values at each site with data indicating potential pollution effects. In addition, relationships between the index values and several habitat variables were examined to ensure that the index was measuring the pollution gradient as intended, rather than the habitat gradients that affect species distributions.

Correlations between index values and the two pollution indicator variables were used as an internal form of validation. Since the index was developed from a linear combination of the two variables, it was necessary to ensure that it adequately reflected habitat alteration.

Relationships between the index values and seven habitat variables were examined to ensure that the index was measuring the intended pollution gradient rather than one or more of the habitat variables. The habitat variables included sediment grain size composition, total organic carbon, water depth, longitude, latitude, and time. Time was included to determine whether consistent inter-annual differences in index values existed due to climate (e.g., El Nino or La Nina) or other effects.

(ii) Comparison with SFBIBI Assessment Results

San Francisco Assessment Method (SFBIBI) assessments were available for 225 samples and San Francisco Benthic Response Index (SFBRI) values for 165. Both measures were available for 142 samples. For these 142 samples, the SFBIBI and SFBRI assessment results were compared using two types of contingency tables. The first compared the coincidence of samples classified as disturbed and undisturbed. The second compared the severity of disturbance indicated by SFBRI response levels with the number of indicator values outside SFBIBI reference ranges. A Mantel-Haenszel chi-square test (Mantel and Haenszel 1959) was used to test for linear associations between results for the two indices.

Results

(i) Development of a Benthic Response Index for San Francisco Bay

1. Assemble Data

Benthic infauna data from 1,153 samples collected for 10 projects were assembled into the project database (Table 1). Sediment contaminant concentration data were available for 277 of these samples and amphipod toxicity data for 197 (Table 2). Both types of data were available for 193 samples.

2. Identify Habitats with Distinct Natural Assemblages

Sequential analysis of the splits on the cluster analysis dendrogram yielded two habitat-related benthic infaunal assemblages in San Francisco Bay (Figures 2 and 3, Table 3). Eighty-four (97.7%) of the 86 samples selected for cluster analysis classified correctly when divided at latitude 37.9° (37°54') into northern and southern assemblages (Table 3). Latitude was the only habitat variable that varied across the dendrogram split (Figure 4) with a highly significant ($p < 0.0001$) Mann-Whitney U-Test for difference in median. The other habitat variables overlapped substantially (Figures 5 through 8) and were not significantly different. Only two assemblages were identified because only the first split met our criteria for habitat differentiation.

Substantial differences were found in species composition between the northern and southern assemblages (Figure 9). Substantial differences also were observed in diversity (Figure 10) and total abundance (Figure 11).

3. Identify the Pollution Vector in Ordination Space

The results of the canonical correlation analysis on the 167 samples selected for BRI development are presented in Table 4, which shows the correlations between the first two ordination axes and the two indicators of environmental pollution. These correlations were used to locate the overall pollution vector in the multivariate ordination space (see Appendix A).

4. Identify the Position of Each Species on the Pollution Vector

The optimization procedure resulted in abundance transformation exponents (f) of 0.33 for both habitats (Table 5), the same exponent that was used for all five southern California BRI habitats. Examples of the relationship between p_i values, species abundance distributions, and the index pollution gradient are presented in Figures 12 and 13. Peak abundances of the species with the highest p_i values in the south Bay habitat occur at higher index values (Figure 12). Peak abundances for species with lower p_i values occur at lower index values (Figure 13).

The list of 17 species with the 10 highest pollution tolerance scores in both habitats included 10 annelids, 5 arthropods, and 2 molluscs (Table 6). All but five occurred in both habitats (Table 6). The taxa with the two highest p_i values in the southern San Francisco Bay habitat, *Capitella capitata* and Oligochaeta, were selected previously by Thompson and Lowe (in press) as indicators of contamination in San Francisco Bay. The most pollution-indicative species, *Capitella capitata*, is well known as an indicator of organic pollution (Grassle and Grassle 1984). *Streblospio benedicti*, another species often associated with disturbance and pollution that was also selected by Thompson and Lowe (in press), also had a pollution tolerance score towards the polluted end. This species had the 8^{th} highest pollution tolerance score in the southern habitat.

The list of 20 species with the 10 lowest pollution tolerance scores was equally diverse. It included 3 ectoprocts and a cnidarian as well as 5 molluscs, 2 arthropods, and 9 annelids (Table 7). Of these species, five species belonged to two genera; three species were from the polychaete genus *Glycera* and two species were from the bivalave genus *Tellina*. Only three species occurred in both habitats; the polychaete worm *Glycera americana*, the Ectoproct *Electra* sp, and the barnacle *Balanus improvisus*.

5. Develop Assessment Thresholds

Assessment thresholds were selected for the index based on changes in biodiversity along the pollution gradient defined by the index values. The portion of the gradient occupied by each species in the northern and southern habitats is presented in Figures 14 and 15 and summarized by the number of species curves in Figures 16 and 17. At the unpolluted (reference) end of the pollution gradients, species appeared and few, if any, dropped out. As a result, the number of potential species increased rapidly. Further along the gradient, the number of species dropping out increased until it equaled the number of species entering, and the net number of potential species stabilized. Eventually, the number of species dropping out exceeded the number of species entering and the number of potential species declined.

Threshold values were established using the number of species curves (Figures 16 and 17). Using these curves, reference thresholds were established at the points where the number of species fell 5% below the peak net number of species; at index values of 19 and 44 for the northern and southern habitats, respectively. Two outlier values were eliminated before identification of the peak in the northern habitat (Figure 16). An arbitrary value of 5% was selected for the following reasons: (1) The southern peak is somewhat flat, making it difficult to identify the point at which the peak occurs, but is followed by a definite region of decline. Thus, 5% below the peak is a better defined point than the peak. (2) The threshold is appropriately placed where net species loss begins to occur, which would be a small amount (we chose 5%) past the peak. (3) Choosing 5% allows for some error in our analyses that might lead to too low of a reference threshold value. Although a single set of thresholds would have been preferred for both habitats, sites in the northern habitat supported fewer species than the southern habitat (Figures 10 and 11). Thus, the decision was made not to combine the two curves to yield an average, which was done for the southern California BRI.

Three more thresholds were defined at the points where 25%, 50%, and 80% of the biodiversity of the reference pool was lost (Table 8). The 25% threshold was defined as the index value where the potential number of species drops to 25% below the number of species ranges that cover the reference threshold. Thus, the basis of the 25% is the number of species that have appeared and not yet dropped out at the reference threshold. The 50% and 80% biodiversity loss thresholds were calculated in a similar fashion.

6. Validate Index Values and Pollution Tolerance Scores

The correlation coefficients between the index and the mean ERM quotients and amphipod mortality were 0.54 and 0.61, respectively, for the northern habitat, and 0.60 and 0.40, respectively, for the southern bay habitat. Both correlations were statistically significant at $p < 0.0001$. The indicator variables explained about half the variation in the index in the northern habitat and just over a third in the south.

Habitat variables did not consistently covary with index values and it was not possible to predict index values from habitat variable values (Figures 18 and 19). In general, associations were weaker in the southern habitat. Most of the samples at the polluted end of the index gradient occurred at shallow depths in both habitats and at low salinity in the northern habitat. However, samples at similar depths and salinity also occurred at the unpolluted end of the gradient. Total organic carbon (TOC) increased along the pollution gradient defined by the index, but a few data points at the polluted end of the gradient contributed disproportionately to the relationship. Sediment grain size distribution (% fines) does not follow the same pattern as TOC, indicating that the increasing TOC is probably from anthropogenic sources rather than organic material naturally adhering to the larger surface area of smaller sized sediments (Newell 1979).

(ii) Comparison with SFBIBI Assessments

Assessments by the two indices were in agreement for 67.6% of the 142 samples for which assessments by both were available (Table 9). Of these samples, 12.0% were assessed as clearly disturbed and 55.6% as undisturbed. The SFBIBI assessed only five (3.5%) of the samples for which assessments disagreed as clearly disturbed, while the SFBRI placed the other 41 (28.9%) in this category. The Mantel-Haenszel chi-square test indicated an association between the results for the two indices was highly significant ($p < 0.001$).

The association between SFBRI and SFBIBI disturbance level assessments was also highly significant ($p < 0.0001$; Table 10). Overall, the SFBRI assessed samples at higher levels of disturbance than the SFBIBI.

Discussion

(i) Development of a Benthic Response Index for San Francisco Bay

A measure of disturbance was developed for San Francisco Bay benthic communities using a BRI approach. There were several indications of success. High pollution tolerance scores were assigned to the polychaete worms *Capitella capitata* and *Streblospio benedicti*, oligochaete worms, and the bivalve *Gemma gemma*, which are widely known as opportunistic, pollution tolerant organisms (Table 6, Grassle and Grassle 1984, Weisberg *et al.* 1997, Llanso *et al.* 2002). Thompson and Lowe (in press) recently selected the first three of these taxa as indicators of contamination in San Francisco Bay. In addition, 12 of the 17 species with the 10 highest pollution tolerance scores occurred in both the north and south habitats while only three of 20

species with the lowest scores occurred in both habitats. Since pollution-sensitive organisms are well adapted to specific habitats, overlap of species is not expected. On the other hand, pollution-tolerant organisms are known to occur across a broad range of habitats, but only when environmental conditions are sufficiently degraded that the sensitive resident organisms can no longer compete.
The assemblages identified using cluster analysis corresponded well with the assemblages identified previously in San Francisco Bay (Thompson *et al.* 2000). There were substantial differences between the northern and southern San Francisco Bay assemblages in abundance (Figure 11), diversity (Figure 10) and species composition (Figure 9). The nature of the relationships between biological community measures and the index pollution gradient also differed between the two assemblages (Figures 20 and 21).

The BRI in San Francisco Bay should be considered preliminary for four reasons. First, insufficient data were available to validate the BRI independently. Accuracy of an assessment measure is confirmed through successful application to sites where benthic condition is known or predictable but are not used in development of the measure. Typically, this is accomplished through a parsing of the original data set, which was not possible due to sample size. Although our database contained data for 1,153 samples, only 167 of them had the accompanying sediment contaminant concentration and amphipod toxicity data necessary for BRI development. In addition, many of the data used to develop the BRI were from repeated sampling at the same site, which further limits our ability to adequately assess the range of conditions encountered in San Francisco Bay.

Second, insufficient data from highly polluted sites were available for full index development. Samples from significantly disturbed sites are required to capture the entire pollution gradient. Only two samples in the northern San Francisco Bay habitat and three samples in the south had median ERM quotients greater than 0.5, which is not substantially impacted on a national scale (Long *et al.* 1998, Long and MacDonald 1998).

Third, relationships between the BRI and pollution indicator variables were weak, though similar to those during BRI development in southern California bays. Pearson correlation coefficients with mean ERM quotients in northern and southern San Francisco bay habitats were 0.54 and 0.60, respectively, while they were 0.52 and 0.65 for the northern and southern California bays. The correlation coefficients with amphipod toxicity were 0.61 and 0.40 in the northern and southern San Francisco Bay habitats, and 0.72 and 0.65 for the northern and southern California bays.

There are several possible reasons for these weak relationships. Two different amphipod species, *Eohaustorius estuarius* and *Ampelisca abdita* were used in the tests; these species differ in their responses to several contaminants, increasing variability about the pollution vector defined in the ordination space. In addition, gear types varied with different sampling programs, all data were included because of the few samples where benthic data were collected synoptically with pollution indicator (sediment contaminants and amphipod toxicity) data. Gear differences may have added variability to the ordination space in which the pollution vector was defined. Another reason may be our choice of index period. Many benthic indices factor out seasonal variation in benthic abundances by restricting applicability to summer, the season with the most stable climate and the most stable benthic populations. After examining the data, we chose not to impose an index period in order to improve definition of the pollution vector because the most contaminated samples were collected in winter or spring. It may be possible to improve the relationship of the BRI with other indicators of impact by imposing an index period, especially if summer data from directed sampling of highly polluted sites are available.

Fourth, the relationships between the BRI and habitat measures, albeit statistically weak, are stronger in San Francisco than for previous applications of this approach. Ideally, benthic indices are highly correlated with pollution indicators such as sediment contaminants and amphipod toxicity, and uncorrelated with habitat measures such as salinity, water depth, sediment grain size distribution, and total organic carbon. Ideally, the index is predictive of disturbance and independent of habitat. In San Francisco Bay, the apparent relationships are driven by few highly contaminated samples collected at very shallow, very muddy sites with high organic carbon. Data from directed sampling at highly polluted sites with a range of habitat characteristics would likely distinguish between pollution and habitat effects.

(ii) Comparison with SFBIBI Assessments

The level of agreement between assessments of undisturbed and clearly disturbed samples by the San Francisco BRI and IBI was highly significant. Where the two indices diverged in San Francisco Bay, the BRI was more conservative than the IBI, classifying as clearly disturbed 41 of 46 samples where assessments disagreed. The BRI sets a lower threshold for identifying disturbance. Unfortunately, no independent measure is available to indicate which threshold is closer to "truth."

The level of agreement between the indices was within the range for classification efficiencies achieved during the development of benthic indices elsewhere (Engle *et al.* 1994, Weisberg *et al.* 1997, Engle and Summers 1999, Van Dolah *et al.* 1999, Paul *et al.* 2001, Llanso *et al.* 2002). Although the indices were developed using different approaches, the level of agreement is expected because most of the development data were the same. Only one previous study (Ranasinghe *et al.* 2002) has contrasted assessment results by estuarine benthic indices developed using different approaches on the same set of samples. In Chesapeake Bay, they compared assessment results for an IBI and an index based on discriminant analysis, which is an approach in the same multivariate analysis family as the BRI, but emphasizing community measures rather than species abundances. In that study, assessments by the Chesapeake Benthic Index of Biotic Integrity and the U.S. EPA's Environmental Monitoring and Assessment Program's Virginian Province Benthic Index agreed for 81.3% of 294 samples. This result was about 14% higher than the level of agreement in our study.

ADAPTATION OF THE SAN FRANSISCO BAY IBI FOR SOUTHERN CALIFORNIA BAYS AND COMPARISON WITH BRI ASSESSMENTS

Methods

(i) Data

Benthic macrofaunal data

Benthic macrofaunal data assembled by SCCWRP from southern California bays and estuaries were used to develop an Index of Biotic Integrity (IBI) and compare assessment results to the Benthic Response Index or BRI (Smith *et al.* 2003). Methods of sample collection and analysis, including taxonomy standards, are detailed in the reports cited below. The data were collected from numerous bays and harbors between Pt. Conception and Mexico (listed on Tables 11, 12A and 12B) as follows:

- Bight'98 Regional Monitoring Program; 96 samples in 1998 (Ranasinghe *et al.* 2003)
- West EMAP, 23 samples in 1999 (Heitmuller *et al.* 1999)
- San Diego Bay TMDL Study, 36 samples in 2004 (Southern California Coastal Water Research Project and Space and Naval Warfare Systems Center San Diego 2004).

Data from samples collected by the Bay Protection and Toxic Clean-up Program (BPTCP) were not used because sampling gear and sieve sizes were different. Additionally, samples classified as North-South Overlap assemblage sites by Smith *et al.* (2003) were omitted; only North and South assemblage sites were used so as to provide the most robust analysis possible.

A total of 155 samples were used, which included 60 from the North assemblage and 95 from the South assemblage (Tables 11, 12A and 12B).

Sediment Data

Sediments were sampled at the same time that the benthic samples were collected. Sediment grain-size, organic content, and a set of sediment contaminants (trace metals, trace organics, pesticides) were sampled and measured. The details of sampling and measurement are presented elsewhere (Heitmuller *et al.* 1999, Ranasinghe *et al.* 2003, Southern California Coastal Water Research Project and Space and Naval Warfare Systems Center San Diego 2004).

One of the primary uses of benthic bioassessments is to evaluate sediment conditions. Therefore, benthic bioassessments must be developed that adequately reflect sediment conditions, especially sediment contamination. Since benthos are known to respond to a variety of sediment factors, an understanding of their relative response to major sediment variables is important. Percent fine sediments (<63 um), percent total organic carbon (TOC), and the mean ERM quotient (mERMq), a sediment contamination "index" were used to identify, select, and test the benthic indicators. The mERMq is described in detail in Long *et al.* (1998). Briefly, effects range-median (ERM) sediment quality guidelines that were "frequently" associated with biological effects (Long *et al.*,

1995) were used to calculate mERMq. Concentrations of 24 contaminants for which ERM values exist, were used. These included eight trace metals (Ag, As, Cd, Cr, Cu, Hg, Pb, Zn), 13 PAH compounds (acenapthene, acenaphthylene, anthracene, benzo[a]pyrene, benzo[a] anthracene, chrysene, dibenz[a,h]anthracene, fluoranthene fluorene, 2-methylnaphthalene, naphthalene, phenanthrene, pyrene), p,p'-DDE, total DDTs, and total PCBs. However, all contaminants were not measured at some sites. Mean ERM quotients computed using these components have been used in previous studies of benthic impacts (Carr *et al.* 1996, Hyland *et al.* 1999).

Data analyses were conducted using the Statistical Analysis System (SAS, 1996). Spearman's rank correlations and multiple regression analyses were used to evaluate relationships between benthic indicators and abiotic variables. In particular, we evaluated the proportion of variance in indicator values accounted for by sediment contamination (mERMq) when percent fines and TOC were included in the analysis. Since the expected responses of the indicators to the abiotic variables were curvilinear, the data were transformed prior to analysis (natural log or arcsine). The regression model that included the combination of transformed and/or untransformed abiotic variables (independent variables) that provided the highest R^2 value for each assessment indicator (dependent variable) was used. These analyses were used only to evaluate the relative contributions of selected abiotic variables to indicator variation, not for predictions of indicator responses. The Wilcoxon 2-sample test with ranked data was used to compare samples statistically.

(ii) Assessment Methods

The identification of benthic assemblages is an important step in bioassessments. Since habitat conditions and species composition within assemblages are relatively homogeneous, the development of assessments for each assemblage minimizes the variability in benthic responses to large differences in salinity or sediment type found in different assemblages. Classification and ordination analysis conducted by Southern California Coastal Water Research Program (SCCWRP) (unpublished) using the southern California Bays and Estuaries data was used as the basis for the assignment of samples to two benthic assemblages: a North assemblage included samples north of Dana Point, and a South assemblage included samples south of Dana Pt. A North-South Overlap sub-assemblage was also identified, but those samples were not used in this assessment.

Two macrobenthic assessment methods will be compared:

- The San Francisco Estuary IBI method (Thompson and Lowe In press)
- The Southern California BRI method (Smith *et al.* 2003)

The IBI method was applied to the Southern California Bays and Estuaries in five steps, which are described in detail in the following section:

(i) Identification of benthic assessment indicators
(ii) Evaluation and selection of indicators
(iii) Identification of reference site and calculation of indicator reference range values

(iv) Assessment of benthic test samples
(v) Evaluation of Results

Results

(i) Identification of Benthic Assessment Indicators

Candidate benthic assessment indicators were identified from the literature (used in other estuarine assessments) including those used in San Francisco Estuary, and from analysis of southern California data to determine which indicators responded best to sediment contamination. The latter step included:

- Identification of the most common and abundant taxa. However, known sensitive taxa that may only be present in a few contaminated samples were considered.

- Spearman's rank correlations between the candidate indicator and mERMq

- Identification of indicator as pollution-tolerant or pollution-sensitive in the literature, including the BRI analysis pollution tolerant scores

Twenty-two candidate benthic assessment indicators were identified for further testing (Table 13). Consistent with principles for IBI development (Karr and Chu 1999), they include community metrics, higher taxa, and indicator species that are both sensitive to and tolerant of sediment contamination.

Four of the candidate indicators were among the 10 lowest pollution tolerant scores (sensitive) taxa, and one was among the 10 highest pollution tolerant scores (tolerant) as determined by Smith *et al.* (2003). In the North assemblage, 12 of the indicators were significantly correlated with mERmq, but mERMq accounted for the generally small proportions of indicator variance in multiple regression analysis (see next section). Mean ERM quotients accounted for significant portions of variance only for abundances of the polychaete family Capitellidae and *Capitella capitata.* In the South assemblage, most of the indicators were significantly correlated with mERMq, and mERMq accounted for significant portions of variation for five of the indicators: molluscan taxa, Dorvillidae, *Dorvillea (S.)* sp., *Notomastus* sp., and *Tagelus subteres* (Table 13).

(ii) Evaluation and Selection of Indicators

Ideally, benthic assessment indicators should conform to current conceptual models of benthic response to contamination and should demonstrate a measurable response to sediment contamination. Responses to sediment-type and organic content should be minimal compared to responses to contamination. All candidate indicators were evaluated to see whether they conformed to this ideal.

First, plots of the candidate indicators versus mERMq were examined for an observable response to sediment contamination. Community-level indicators (total taxa, abundance) should respond to sediment contamination gradients in a nonlinear manner per the Pearson and Rosenberg (1978) model, where benthic indicator response to moderate contamination may result in elevated values compared to reference conditions. Severe sediment contamination results in greatly reduced indicator values. Higher taxa and tolerant species may also respond in a nonlinear manner, but sensitive taxa may respond in a more linear manner (Thompson and Lowe In press). However, if a full range of sediment contamination (uncontaminated to severely contaminated) is not sampled, the response plots may reflect only part of the benthic response. Plots for all 22 candidate indicators are not shown, but plots of selected indicators versus mERMq are shown on Figure 22. Various forms of response to mERMq can be seen. None show the classic Pearson-Rosenberg response. Total taxa and molluscan taxa exhibited monotonic responses to sediment contamination, and total abundance, *Dorvillea* (*S.*) sp. and *Spiophanes duplex* showed some indication of curvilinear response in one or the other assemblage. The indicator responses in these plots could be partially due to indicator responses to sediment-type and TOC, known to covary with sediment contamination.

Multiple regression analysis was conducted to evaluate the relationships between candidate benthic indicators and percent fines, TOC, and mERMq (independent variables). These analyses showed that sediment contamination (mERMq) accounted for the significant portions of indicator variation for only a few indicators (Table 13). Results of the multiple regression analysis for the selected assessment indicators are shown on Table 14. The multiple regression models generally accounted for low proportions of variance in the indicators, less than 37.5%; but all regression models were significant ($p<0.05$). This analysis also showed the relative response by the candidate indicator to sediment grain-size, TOC, and mERMq. In the North assemblage, sediment grain size accounted for the largest proportion of variance in the models for all indicators except for *Dorvillea (S.)* sp., where TOC accounted for most of the variance. Mean ERM quotient accounted for 7.5 – 26.7% of model variation, and was not a significant component of any of the models. However, mERMq was significantly correlated with all indicators ($p<0.05$, Table 13). These indicators responded more to sediment grain size and TOC than to contamination. Therefore, assessments using these indicators will demonstrate a response to both sediment type and contamination. In the South assemblage, TOC accounted for the largest proportion of model variation for total taxa and total abundance, but mERMq accounted for the largest proportion of variance for molluscan taxa, and was a significant component. Assessments using these indicators will demonstrate responses to both TOC and contamination.

Seven indicators from each assemblage were selected for further optimization testing based on their widespread frequency and abundance, known contaminant tolerance, response to mERMq, and ability to provide adequate reference ranges for assessment (Table 13). Various combinations of these indicators were tested to determine which combination optimized the correlation between assessment values (AV, described below) and mERMq.

From the above evaluations and analyses, four benthic assessment indicators were selected for use in the North assemblage, and three were selected for use in the South assemblage (Table 14).

(iii) Identification of Reference Site and Indicator Reference Range Values

A reference sample screening process was used similar to that used in San Francisco Estuary. All 155 sites were screened using four criteria that generally define reference sites:

- No sediment toxicity
- Composed mostly of contaminant sensitive taxa. This was accomplished using a scaled ratio of sensitive:tolerant taxa (Thompson and Lowe In press).
- Presence of amphipods (DeWitt *et al.* 1989)
- Absence of *Capitella capitata*

These criteria were slightly different than those used in San Francisco Estuary assessments, where oligochates were used instead of *C. capitata*. Oligochates were rare in the southern California samples, probably because a 1.0 mm sieve size was used. In the San Francisco Estuary, a sample was considered to be a reference site if it was not toxic and none, or any one of the three other criteria was exceeded. However, in southern California, applying that selection procedure resulted in overly liberal reference site selection. A total of 100 samples were identified as reference, and many of those samples had mostly tolerant taxa (sensitive:tolerant ratio up to 0.88), or elevated abundances of *C. capitata.* It also included many samples with relatively high mERMq values, up 0.713 in the North, and up to 0.423 in the South.

A modified version of the reference selection procedure was used for the southern California assessments. Samples were not considered to be reference sites if <u>any</u> of reference screening criteria were exceeded. This approach yielded 34 reference samples. Nineteen samples from the North assemblage in Los Angeles, Long Beach, and King Harbors, and Anaheim Bay; and fifteen samples from the South assemblage in San Diego and Mission Bays were identified as reference samples (Table 11). These southern California reference samples were not toxic, were composed of mostly sensitive taxa, had amphipods, and no *C. capitata*. Although not used in the reference sample selection procedure, the mERMq values at these reference sites ranged up to 0.477 in the North assemblage and up to 0.159 in the South.

The minimum and maximum value of the benthic assessment indicators in the reference samples from each assemblage is shown on Table 15. These values will be used to assess whether the indicator values in the remaining non-reference samples are within or outside of these reference ranges.

(iv) Assessment of Benthic Samples

The San Francisco Bay IBI method uses a weight of evidence approach. The value of each benthic assessment indicator in each sample was compared to the reference range value for the respective assemblage (Table 15). If the test sample value was above or below the reference range it was considered to be a ‘hit’. According to current benthic response theory, some benthic indicator values may be above the reference range in moderately impacted areas. The sum of ‘hits’ for all the assessment indicators in each sample provides an assessment value (AV) that reflects the degree of impact in the sample. A sample was considered to be impacted if two or more of the reference ranges were exceeded. If none, or only one reference range was exceeded, it was considered unimpacted, providing benefit of doubt for samples that only exceed one reference range. Higher numbers of reference ranges exceedances provide evidence for more severe impacts.

The San Francisco Estuary IBI assessment method was applied two ways: (1) using the San Francisco Estuary assessment method *per se*, and using a modified procedure. Application *per se* did not result in meaningful results and was modified by changing the reference site selection criteria and using a different set of benthic indicators better suited to southern California.

San Francisco Estuary Procedure. Reference ranges from the 100 reference samples identified using one reference criterion exceedance (described above) were used. Non-reference samples were assessed using most of the same benthic assessment indicators used in the San Francisco Estuary polyhaline assemblage assessment (Thompson and Lowe In press). Total taxa, total abundances, molluscan taxa, and *C. capitata* were used; oligochaetes were not used as there were none in the data.

Only five samples from each of the North and South assemblages had samples that exceeded reference ranges (Table 16). This is clearly an unreasonable result. Allowing exceedance of one reference criterion produced very wide reference ranges and most samples fell within those liberal limits.

Modified IBI Procedure. Forty-one samples from the North assemblage and 80 samples from the South assemblage were assessed using the modified assessment procedure. The reference samples listed on Table 11, and reference ranges for the assessment indicators listed on Table 15 were used. Twelve samples from the North assemblage were considered to be to be impacted (AV>1, Table 12A). The impacted samples were collected from nine locations between Santa Barbara and Anaheim Bay. The number of taxa and number of molluscan taxa exceeded reference ranges in many of the samples, but *Dorvillea (S.)* sp. and *S. duplex* only exceeded reference ranges in a few samples. An AV of 2 was the maximum number of indicators that exceeded reference ranges, indicating only "slight" benthic impact. Forty-nine samples from the South assemblage were impacted (Table 12B). The impacted samples were mostly from San Diego Bay, but samples from Mission Bay, Santa Margarita River, and Agua Hedionda Lagoon were also impacted. Molluscan taxa most often exceeded the reference range, but total taxa and total abundances also exceeded reference ranges at many sites. Assessment values ranged between 2 and 3 at the impacted sites, indicating "moderate" to "severe" impacts.

(v) Evaluation of Results

Statistical comparisons of sediment contamination (mERMq) in the reference and impacted samples were used to evaluate whether the assessment results reflected significant differences in sediment contamination (Table 17). Reference samples had significantly lower mERMq values than impacted samples in both assemblages. Therefore, the assessments method correctly reflected elevated sediment contamination. However, percent fine sediments and TOC were also significantly different between reference and impacted samples in the North assemblage. These results are consistent with the multiple regression analysis results, and suggests that, along with sediment contamination, changes in sediment type and TOC may also influence the assessment results. In the South assemblage, both percent fines and TOC were similar between the two assemblages, suggesting that contamination was the primary influence on assessment results.

Increasing AV values also reflected increasing sediment contamination. The AVs shown on Tables 12A and 12B were significantly correlated with mERMq in both assemblages (North: Spearman's r = 0.321, p = 0.40, n=41; South: r = 0.535, p = <0.001, n = 80). This suggests that the degree of impact increases with sediment contamination.

(vi) Comparison to BRI Assessments

The results of the IBI assessments of the southern California bays and estuaries were compared to the results of the Southern California BRI assessments at the same sites. The BRI values used were those presented in the Bight '98 Report (Ranasinghe *et al.* 2003) and calculated by SCCWRP for the West EMAP and San Diego Bay studies.

Reference sites comparison. The BRI method identified a total of 70 reference samples, compared to only 34 identified by IBI. Of the 34 reference sites identified by the southern California IBI assessment described above, 33 (97.1%) were also designated as a reference sites by the BRI method. Site B2228 in San Diego Bay was designated as a reference sample for IBI, but it was assessed as an impact level 1 sample by BRI. One major difference in reference sample designation is that the IBI method does not allow a toxic sample to be a reference sample. The BRI did not consider sediment toxicity in its assessment level designations. Therefore, some reference sites were toxic.

The BRI identified samples with higher mERMq ranges as reference than the IBI. Reference samples identified by the IBI method had mERMq values that ranged up to 0.159 in the South and up to 0.477 in the North. Reference samples identified by the BRI method had mERMq values that ranged up to 0.252 in the South and up to 0.559 in North.

Assessment comparison. Both methods produced results that reflected increasing contaminant concentrations in sediments (Table 18). The proportion of impacted samples (IBI assessment level >1, BRI response level > 0) corresponding to several mERMq ranges was very similar in the North assemblage. However, in the South assemblage the IBI method generally identified fewer samples as impacted within similar mERMq ranges than the BRI method.

The results of the IBI and BRI assessments were also compared using contingency tables. Classification as either impacted or unimpacted showed that the two methods agreed on 41.9% of the samples as unimpacted, and 36.1% of the samples as impacted, for overall agreement on 78.1% of the samples (Table 19). Overall, the BRI designated more samples as impacted than IBI, and designated more samples with low mERMq as impacted than IBI. In the North assemblage, 24.4% of the designations as impacted or unimpacted samples were different between the two methods. Of those, three occurred when IBI designated sites as impacted, and BRI did not. IBI designated seven samples as unimpacted where BRI designated them as impacted. In the South, 27.5 % of the designations were different between the two methods. Of those, only two were where IBI designated sites as impacted, and BRI did not. IBI designated 20 samples as unimpacted where BRI designated them as impacted. Comparisons using more detailed levels of classification, IBI assessment values, and BRI response levels, showed that 57% of the samples agreed in level of classification (Table 20). Agreement was very good within plus or minus one classification level. Although, the classifications used in both tables were significantly different, there was more agreement between the methods than not.

Discussion

Both the IBI and BRI methods effectively distinguished reference and impacted sites. Of those that were assessed, 78.1% of the assessments were the same. However, there were some differences in the designations. The BRI method identified about twice as many sites as reference samples than the IBI method, and identified more sites as impacted than IBI. The reasons for differences in the assessments are not clear. Since there is no standard against which to compare, the "right answer" must be determined by agreement on general principles of assessment application and based on sound science.

The BRI method is more objective and includes information for numerous taxa. More information generally creates more robust indices. However, the BRI method is numerically complex in its formulation and application, and requires considerable expertise to create and calculate the Index. The IBI method is relatively simple in concept and easier to apply. But, identification of indicators requires considerably more subjective judgment than the BRI. The IBI uses fewer indicators and thus may be less robust over a range of samples from varying habitats within an assemblage, and over time.

Very few of the indicators tested exhibited a strong response to sediment contamination. Multiple regression analysis showed that the macrobenthos was only weakly related to sediment contamination in the samples analyzed. Although some taxa exhibited significant correlations with mERMq, when considered together with sediment grain-size and TOC, it was difficult to find benthic indicators that had strong relationships with sediment contamination. Therefore, the IBI assessments may not be as strong as they could be. As noted above, the assessment results cannot be attributed to sediment contamination alone, although contamination was shown to have some influence. Despite the weak relationships with sediment contamination, the IBI method reflected increases in sediment contamination at the impacted sites.

Compared to other areas of the U.S., reference samples from southern California bays and estuaries appear to have higher levels of sediment contamination. In Atlantic and Gulf coast estuaries, the mERMq threshold for moderate or medium benthic impact risks (31-52% of samples) ranged between 0.013 - 0.022, and for high impact risks (55-85% of samples) ranged between 0.036 - 0.098 (Hyland *et al.* 2003). Mean ERM quotient values above 0.20 resulted in marked decreases in number of benthic taxa and arthropod abundance in Florida's Biscayne Bay (Long *et al.* 2002). In San Francisco Bay, reference samples had mERMq values below 0.146 (Thompson and Lowe In press).

CONCLUSIONS

- Distinctly separate approaches have been used to develop benthic assessment tools in southern California and San Francisco Bay.

  In southern California, scientists have developed the Benthic Response Index (BRI), which is a multi-variate approach that uses an abundance-weighted average of the pollution tolerance scores for most of the species that occur at a site. The BRI has successfully been used in offshore regions of southern California and is currently being extended into bays and harbors of that region. In San Francisco Bay, scientists have recently developed an Index of Biological Integrity (IBI), which is a multi-metric approach that uses a set of selected key indicators to distinguish reference from impacted benthic conditions. The BRI method is more objective and includes information for numerous taxa. More information generally creates more robust indices. However, the BRI method is numerically complex in its formulation and application, and requires considerable expertise to create and calculate. The IBI method is relatively simple in concept and easier to apply, but identification of indicators requires considerably more subjective judgment than the BRI.

- Preliminary development of the BRI in San Francisco Bay and the IBI in southern California Bays and Harbors was successful. Benthic indices with thresholds of concern were created, but require further refinement and validation

  The BRI method created a list of pollution tolerance scores for over 200 different species and established five categories of benthic disturbance based on gradations of net species loss. The IBI method identified reference conditions for five indicator metrics in southern California and established a single threshold based on exceedance of reference ranges for at least two of these indicator metrics.

  The application and evaluation of these new assessment methods is limited by several factors. These limitations include: (1) lack of independent data for validation of the index; (2) insufficient data from highly disturbed sites to define the entire range of the impact gradient; and (3) uncertainty in the effect of environmental variables that can affect assemblage distributions regardless of pollution impacts.

- Although the BRI and IBI methods are developed on different premises, they produced reasonably similar results within each region.

  Assessments of the same samples by the BRI and IBI in San Francisco Bay were in agreement for 67.6% of the samples. Similar assessments using the Southern California

BRI and IBI were in agreement for 78.1% of the samples. In both cases, the correlation between results for the two indices was highly significant. In both southern California and in San Francisco Bay, the BRI was more conservative; it classified benthos as disturbed at more sites than the IBI. While there was a relationship with increased sediment contaminant levels, the correlations were generally weak.

LITERATURE CITED

Anderson, B.S., J.W. Hunt, B.M. Phillips and R. Fairey. 2001. Sediment quality in Los Angeles Harbor, USA: A triad assessment. *Environ. Toxicol. Chem.* **20**:359-370.

Bay Area Dischargers Association. 1994. Local effects monitoring program, quality assurance project plan. Bay Area Clean Water Association. Oakland, CA.

Beals, E.W. 1973. Ordination: Mathematical elegance and ecological naivete. *Journal of Ecology* **61**:23-35.

Bergen, M., D.B. Cadien, A. Dalkey, D.E. Montagne, R.W. Smith, J.K. Stull, R.G. Velarde and S.B. Weisberg. 2000. Assessment of benthic infaunal condition on the mainland shelf of Southern California. *Environ. Mon. and Assess.* **64**:421-434.

Bergen, M., S.B. Weisberg, R.W. Smith, D.B. Cadien, A. Dalkey, D.E. Montagne, J.K. Stull, R.G. Velarde and J.A. Ranasinghe. 2001. Relationship between depth, sediment, latitude, and the structure of benthic infaunal assemblages on the mainland shelf of southern California. *Mar. Biol.* **138**:637-647.

Bradfield, G.E. and N.C. Kenkel. 1987. Nonlinear ordination using shortest path adjustment of ecological distances. *Ecology* **68**:750-753.

Bray, J.R. and J.T. Curtis. 1957. An ordination of the upland forest communities of southern Wisconsin. Ecological Monographs 27:325-349. *Ecol. Monogr.* **27**:325-349.

Carr, R.S., E.R. Long, H.L. Windom and D.C. Chapman. 1996. Sediment quality assessment studies of Tampa Bay Florida. *Environ. Toxicol. Chem.* **15**:1281-1231.

Clifford, H.T. and W. Stephenson 1975. <u>An introduction to numerical classification</u>. New York, NY, Academic Press.

Cooley, W.W. and P.R. Lohnes 1971. <u>Multivariate Data Analysis</u>. New York, John Wiley & Sons.

Dauer, D.M., J.A. Ranasinghe and S.B. Weisberg. 2000. Relationships between benthic community condition, water quality, sediment quality, nutrient loads, and land use patterns in Chesapeake Bay. *Estuaries* **23**:80-96.

Department of Water Resources. 1997. Water quality conditions in the Sacramento-San Joaquin Delta during 1994. California Department of Water Resources. Sacramento, CA.

DeWitt, T.H., R.C. Swartz and J.O. Lamberson. 1989. Measuring the acute toxicity of estuarine amphipods. *Environ. Toxicol. Chem.* **8**:1035-1048.

Dillon, W.R. and M. Goldstein 1984. <u>Multivariate Analysis</u>. New York, John Wiley & Sons.

Engle, V.D. and J.K. Summers. 1999. Refinement, validation, and application of a benthic condition index for Northern Gulf of Mexico estuaries. *Estuaries* **22**:624-635.

Engle, V.D., J.K. Summers and G.R. Gaston. 1994. A benthic index of environmental condition of Gulf of Mexico estuaries. *Estuaries* **17**:372-384.

Gittins, R. 1979. Ecological applications of canonical analysis. Multivariate Methods in Ecological Work. L. Orloci, R. Rao and M. Stiteler. Fairland, Maryland, USA, International Co-operative Publishing House. **7:** 309 -535.

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

Grassle, J.P. and J.F. Grassle 1984. The utility of studying the effects of pollutants on single species populations in benthos of mesocosms and coastal ecosystems. Concepts in marine pollution measurements. H.H. White. College Park, MD, Maryland Sea Grant College**:** 621-642.

Hale, S.S., J.F. Paul and J.F. Heltshe. 2004. Watershed landscape indicators of estuarine benthic condition. *Estuaries* **27**:283-295.

Heitmuller, T., W. Benson, F. Mayer, J.K. Summers, J. Moore and G. Craven. 1999. EMAP Western Pilot Coastal Monitoring Quality Assurance Project Plan. U.S. Environmental Protection Agency, Gulf Ecology Division. Gulf Breeze, FL.

Hunt, J.W., B.S. Anderson, B.M. Phillips, R.S. Tjeerdema, K.M. Taberski, C.J. Wilson, H.M. Puckett, M. Stephenson, R. Fairey and J.M. Oakden. 2001. A large-scale categorization of sites in San Francisco Bay, USA, based on the sediment quality triad, toxicity identification evaluations, and gradient studies. *Environ. Toxicol. Chem.* **20**:1252-1265.

Hyland, J.L., W.L. Balthis, V.D. Engle, E.R. Long, J.F. Paul, J.K. Summers and R.F. Van Dolah. 2003. Incidence of stress in benthic communities along the US Atlantic and Gulf of Mexico coasts within different ranges of sediment contamination from chemical mixtures. *Environ. Mon. and Assess.* **81**:149-161.

Hyland, J.L., R.F. Van Dolah and T.R. Snoots. 1999. Predicting stress in benthic communities of southeastern U.S. estuaries in relation to chemical contamination of sediments. *Environ. Toxicol. Chem.* **18**:2557-2564.

Jacobi, M., R. Fairey, C.A. Roberts and E. Landrau. 1998. Chemical and biological measures of sediment quality and tissue bioaccumulation in the north coast region. California State Water Resources Control Board. Sacramento, CA.

Karr, J.R. and E.W. Chu 1999. Restoring life in running waters: Better biological monitoring. Covelo, CA, Island Press.

Lance, G.H. and W.T. Williams. 1967. A general theory of classificatory sorting strategies. I. Hierarchical systems. *Computer Journal* **9**:373-380.

Llanso, R.J., L.C. Scott, J.L. Hyland, D.M. Dauer, D.E. Russell and F.W. Kutz. 2002. An estuarine index of biotic integrity for the mid-Atlantic region of the United States. II. Index Development. *Estuaries* **25**:1231-1242.

Long, E.R., L.J. Field and D.D. MacDonald. 1998. Predicting toxicity in marine sediments with numerical sediment quality guidelines. *Environ. Toxicol. Chem.* **17**:714-727.

Long, E.R., M.J. Hameedi, G.M. Sloan and L.B. Read. 2002. Chemical contamination, toxicity, and benthic community indices in sediments of the Lower Miami River and adjoining portions of Biscayne Bay, Florida. *Estuaries* **25**:

Long, E.R. and D.D. MacDonald. 1998. Recommended uses of empirically derived, sediment quality guidelines for marine and estuarine ecosystems. *Human and Ecological Risk Assessment* **4**:1019-1039.

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

Mantel, N. and W. Haenszel. 1959. Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute* **22**:719-748.

Paul, J.F., K.J. Scott, D.E. Campbell, J.H. Gentile, C.S. Strobel, R.M. Valente, S.B. Weisberg, A.F. Holland and J.A. Ranasinghe. 2001. Developing and applying a benthic index of estuarine condition for the Virginian Biogeographic Province. *Ecological Indicators* **1**:83-99.

Pearson, T.H. and R. Rosenberg. 1978. Macrobenthic succession in relation to organic enrichment and pollution of the marine environment. *Oceanogr. Mar. Biol. Ann. Rev.* **16**:229-311.

Pielou, E.C. 1984. <u>The interpretation of ecological data</u>. New York, John Wiley & Sons.

Ranasinghe, J.A., J.B. Frithsen, F.W. Kutz, J.F. Paul, D.E. Russell, R.A. Batiuk, J.L. Hyland, K.J. Scott and D.M. Dauer. 2002. Application of two indices of benthic community condition in Chesapeake Bay. *Environmetrics* **13**:499-511.

Ranasinghe, J.A., D.E. Montagne, R.W. Smith, T.K. Mikel, S.B. Weisberg, D.B. Cadien, R.G. Velarde and A. Dalkey. 2003. Southern California Bight 1998 Regional Monitoring Program: VII. Benthic Macrofauna. Southern California Coastal Water Research Project. Westminster, CA.

Smith, R.W. 1976. Numerical analysis of ecological survey data, University of Southern California**:** 1 -401.

Smith, R.W., M. Bergen, S.B. Weisberg, D.B. Cadien, A. Dalkey, D.E. Montagne, J.K. Stull and R.G. Velarde. 2001. Benthic response index for assessing infaunal communities on the southern California mainland shelf. *Ecological Applications* **11**:1073-1087.

Smith, R.W., B.B. Bernstein and R.L. Cimberg 1988. Community-environmental relationships in the benthos: Applications of multivariate analytical techniques. Marine Organisms as Indicators. D.F. Soule and G.S. Kleppel. New York, Springer-Verlag**:** 247-326.

Smith, R.W., J.A. Ranasinghe, S.B. Weisberg, D.E. Montagne, D.B. Cadien, T.K. Mikel, R.G. Velarde and A. Dalkey. 2003. Extending the Southern California Benthic Response Index to Assess Benthic Condition in Bays. Southern California Coastal Water Research Program. Westminster, CA.

Southern California Coastal Water Research Project and Space and Naval Warfare Systems Center San Diego. 2004. Sediment assessment study for the mouths of Chollas and Paleta Creek, San Diego. Phase I Draft Report. San Diego Regional Water Quality Control Board, Commander Navy Region Southwest, City of San Diego. San Diego, CA.

Summers, J.K. 2001. Ecological condition of the estuaries of the atlantic and gulf coasts of the United States. *Environ. Toxicol. Chem.* **20**:99-106.

Thompson, B., B.S. Anderson, J.W. Hunt, K.M. Taberski and B.M. Phillips. 1999. Relationships between sediment contamination and toxicity in San Francisco Bay. *Mar. Environ. Res.* **48**:285-395.

Thompson, B. and S. Lowe. In press. Benthos response to contamination in San Francisco Estuary. *Environ. Toxicol. Chem.*

Thompson, B., S. Lowe and M. Kellogg. 2000. Results of the Benthic Pilot Study 1994 - 1997. Part 1. Macrobenthic assemblages of the San Francisco Bay-Delta and their responses to abiotic factors. San Francisco Estuary Regional Monitoring Program for Trace Substances. Richmond, CA. 39.

Thompson, B., S. Lowe and J. Ross. 2002. Benthic Macrofauna. UC Davis. Davis, CA. EPA Grant Number: R826940.

Van Dolah, R.F., J.L. Hyland, A.F. Holland, J.S. Rosen and T.R. Snoots. 1999. A benthic index of biological integrity for assessing habitat quality in estuaries of the southeastern USA. *Mar. Environ. Res.* **48**:269-283.

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

Williamson, M.H. 1978. The ordination of incidence data. *Journal of Ecology* **66**:911-920.

FIGURES

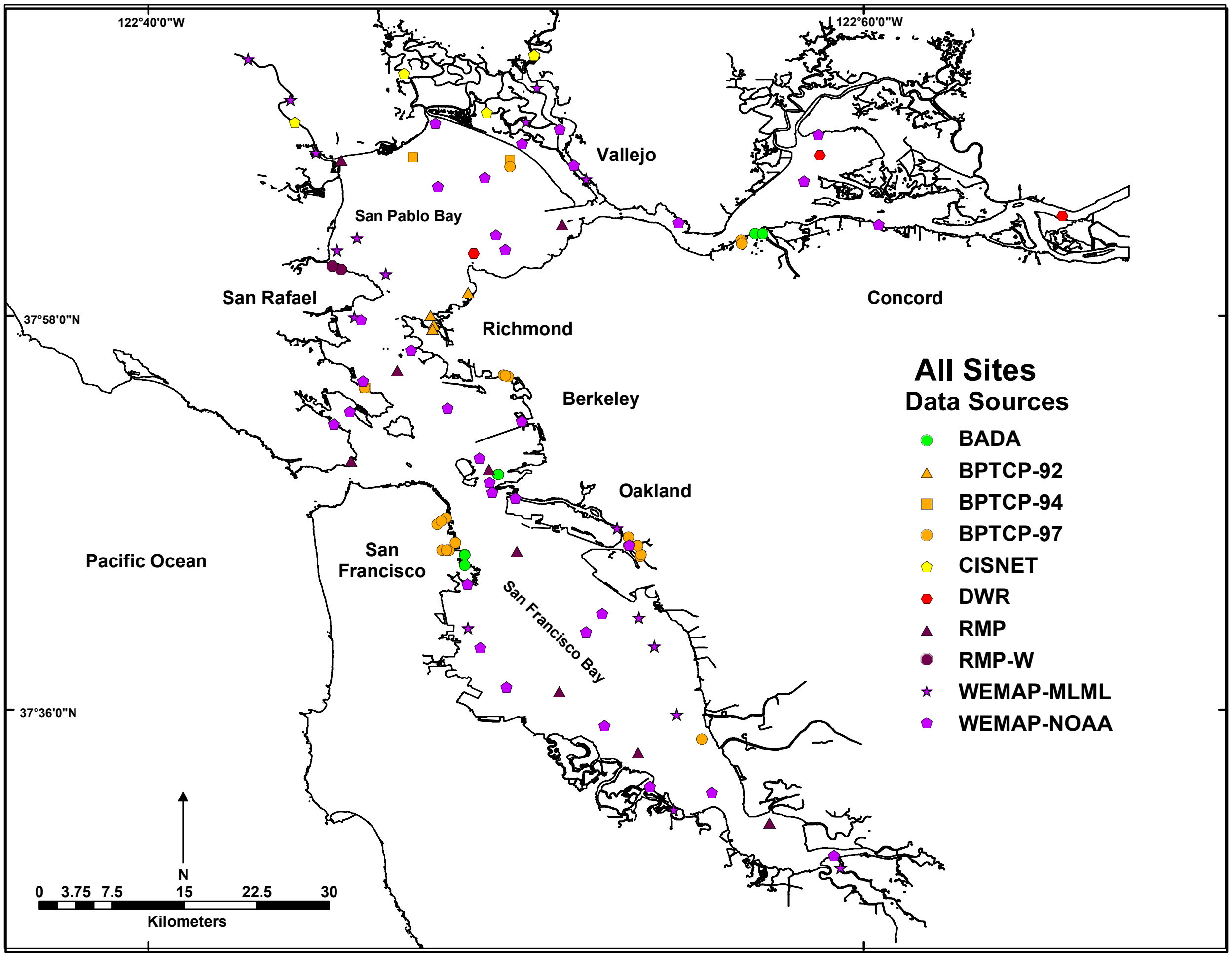

Figure 1. Locations of sampling sites in San Francisco Bay.

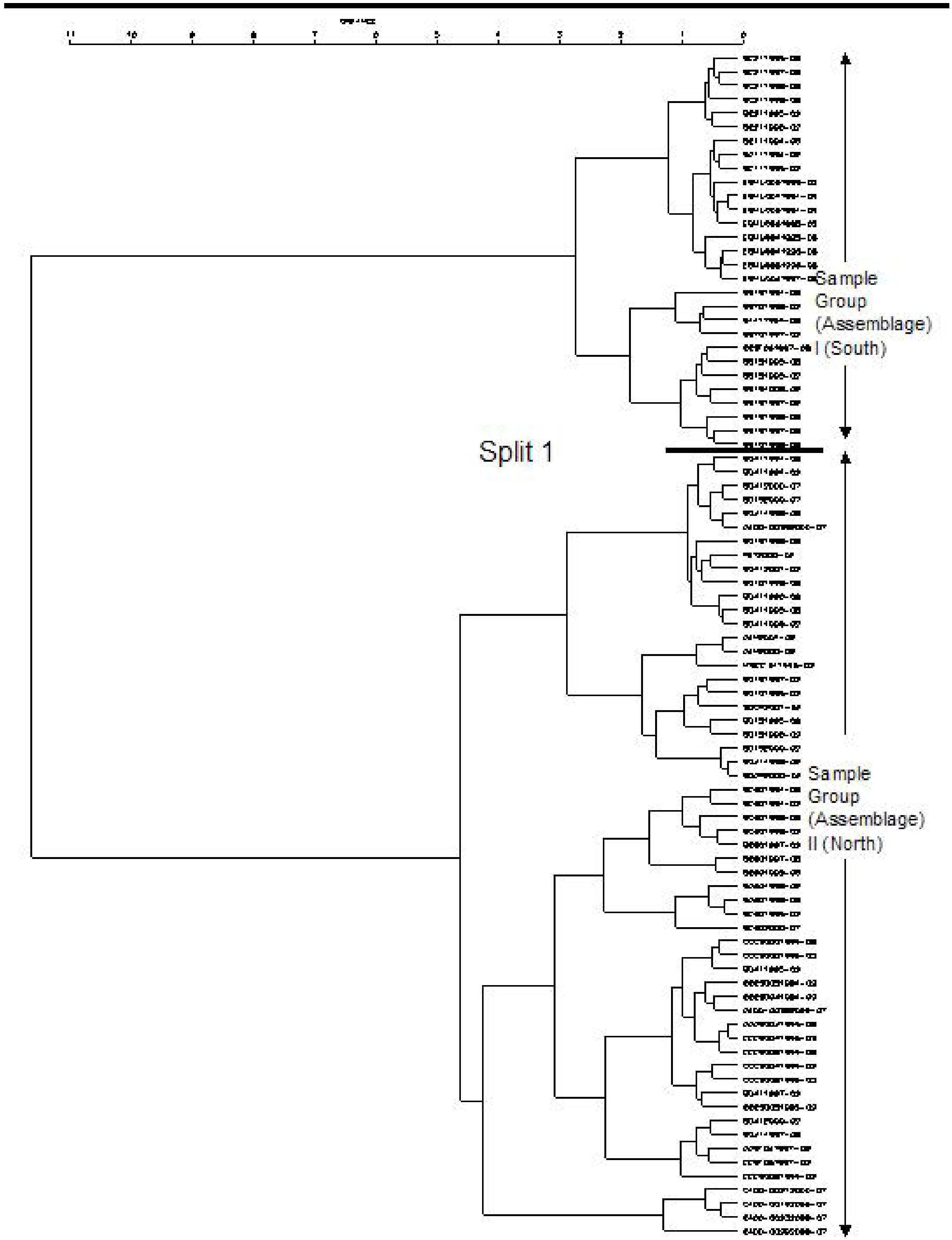

Figure 2. Dendrogram showing assemblages (sample groups) identified by cluster analysis. Split 1 identifies the dendrogram branch point referred to in the text.

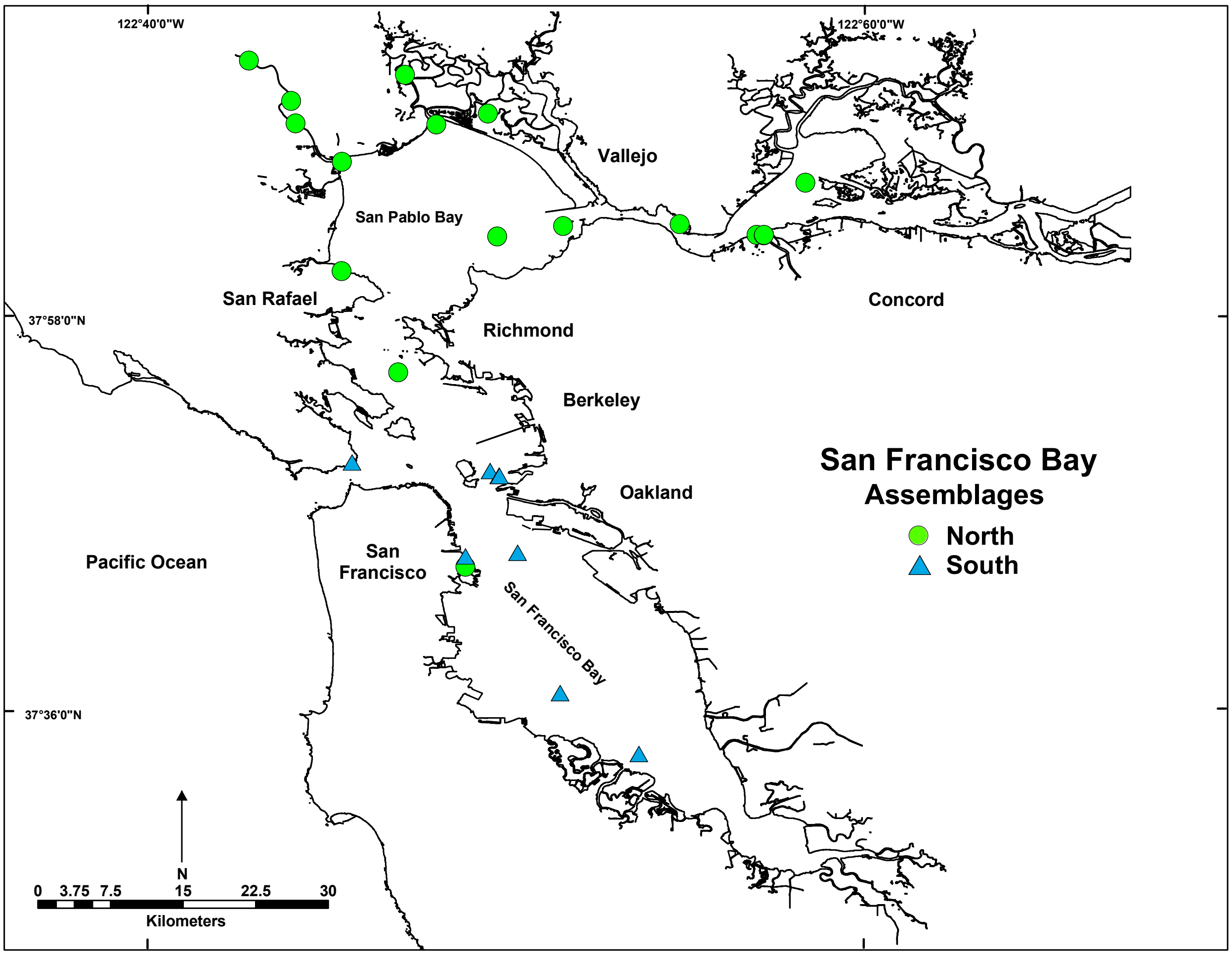

Figure 3. Locations of sites in each assemblage (sample group).

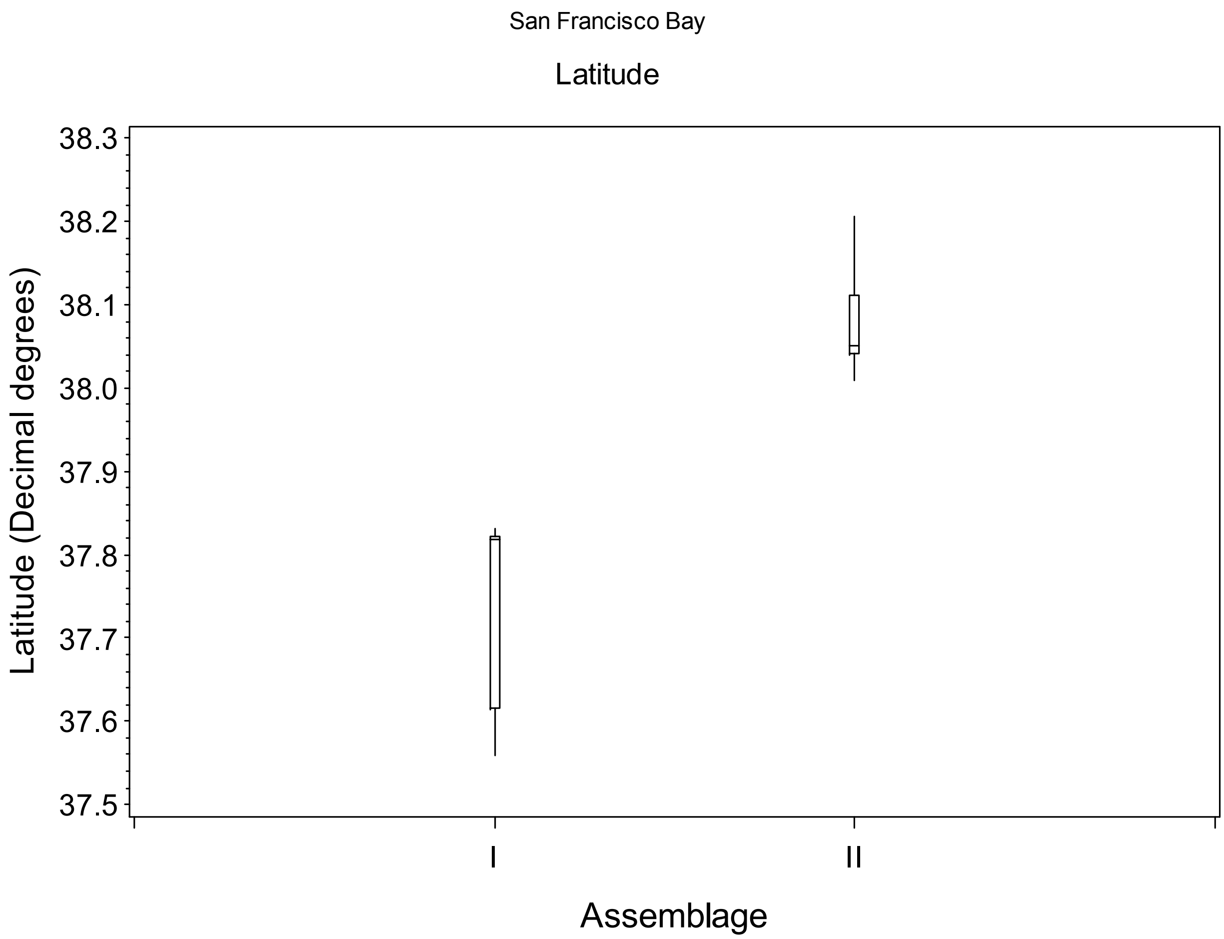

Figure 4. Box plot of latitude distribution for each assemblage (sample group). The bottom and top edges of the box are located at the sample 25th and 75th percentiles. The center horizontal line is the median. The whiskers are drawn from the box to the most extreme point within 1.5 interquartile ranges. An interquartile range is the distance between the 25th and 75th percentiles.

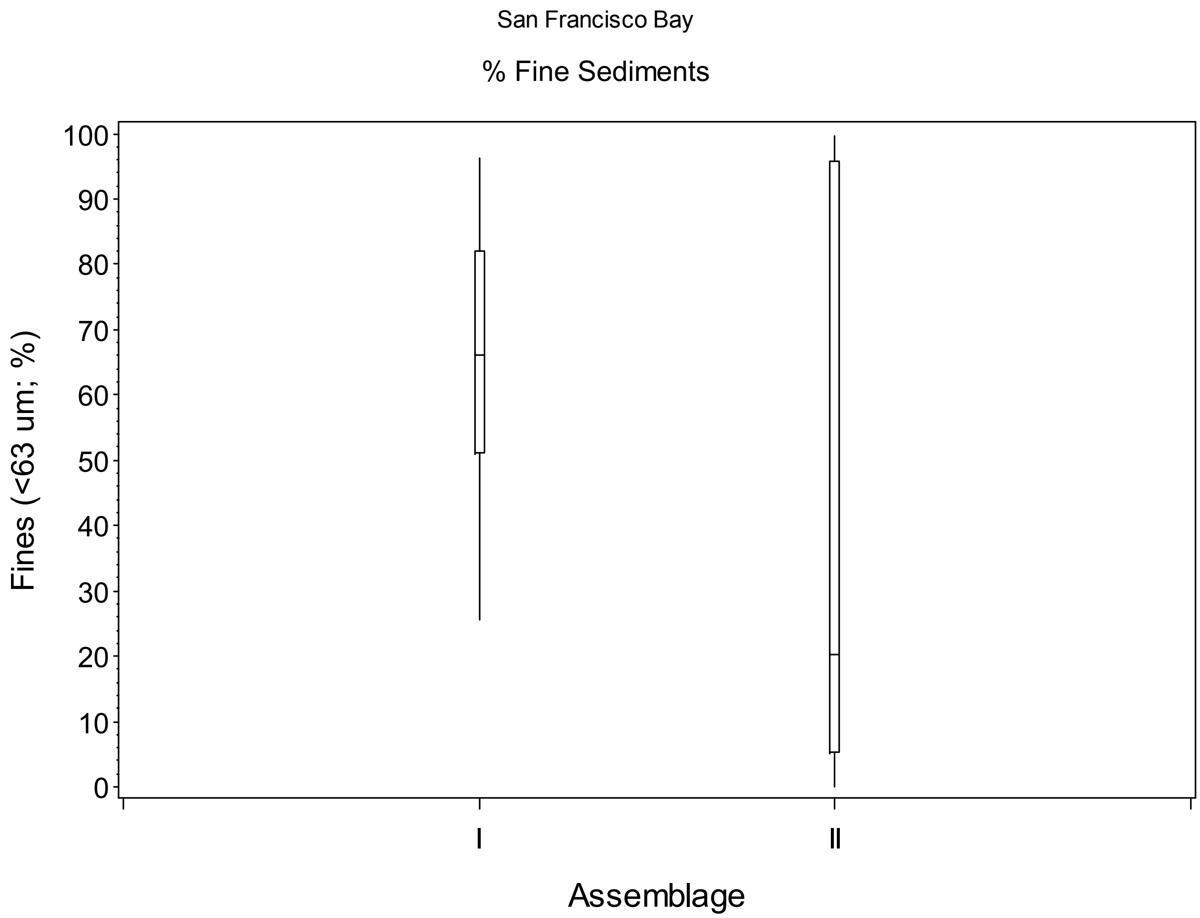

Figure 5. Box plot of fine (< 63μ diameter) sediment distribution for each assemblage (sample group). See Figure 4 for an explanation.

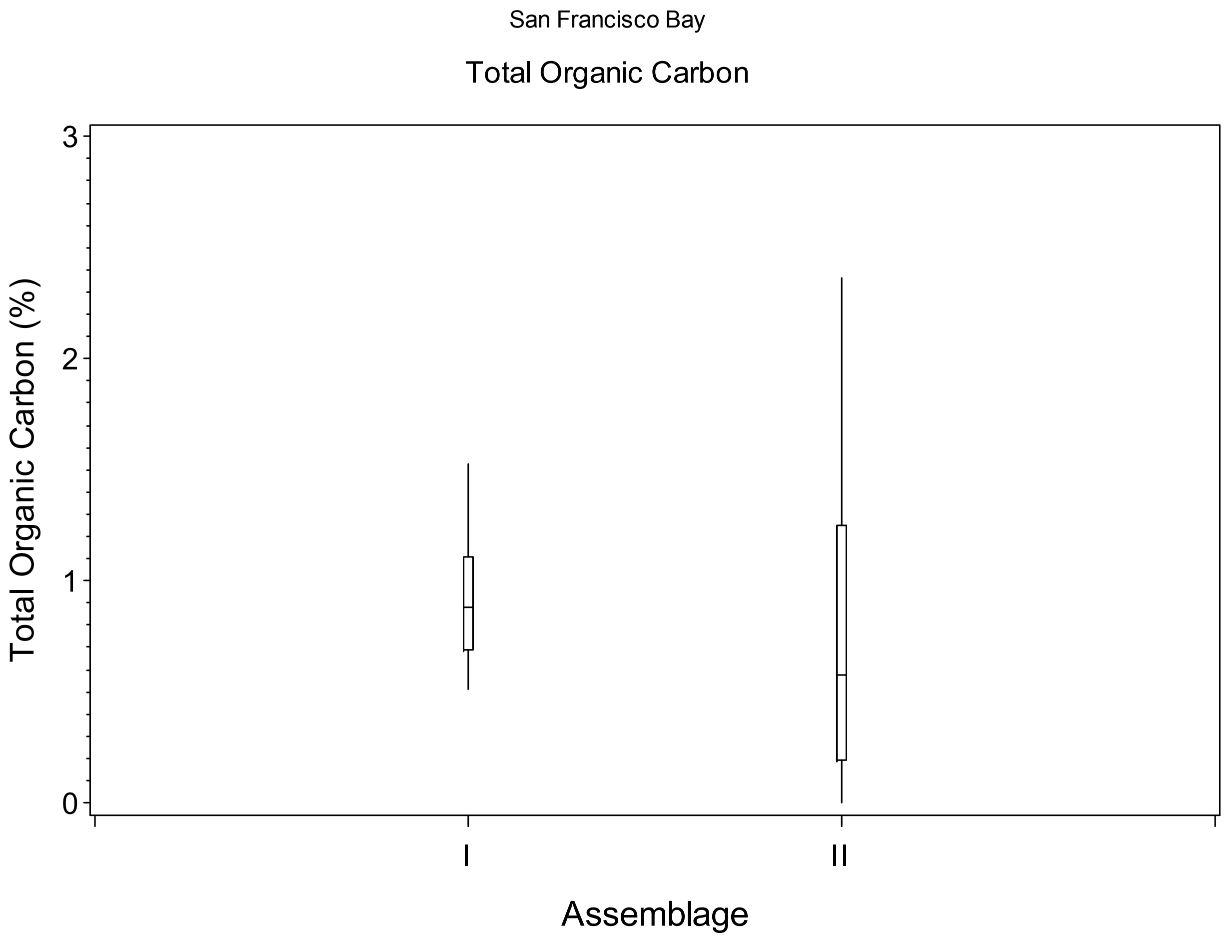

Figure 6. Box plot of total organic carbon content distribution for each assemblage (sample group). See Figure 4 for an explanation.

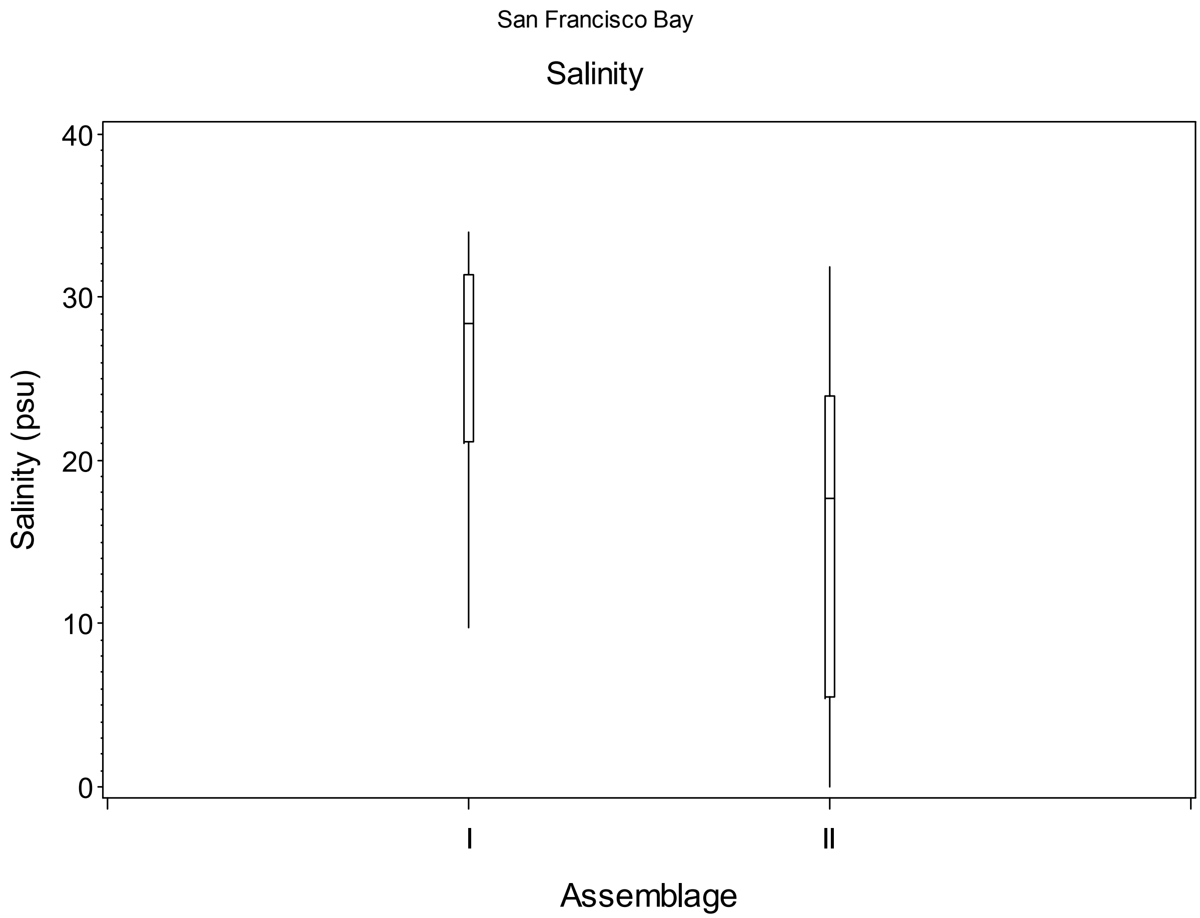

Figure 7. Box plot of salinity distribution for each assemblage (sample group). See Figure 4 for an explanation.

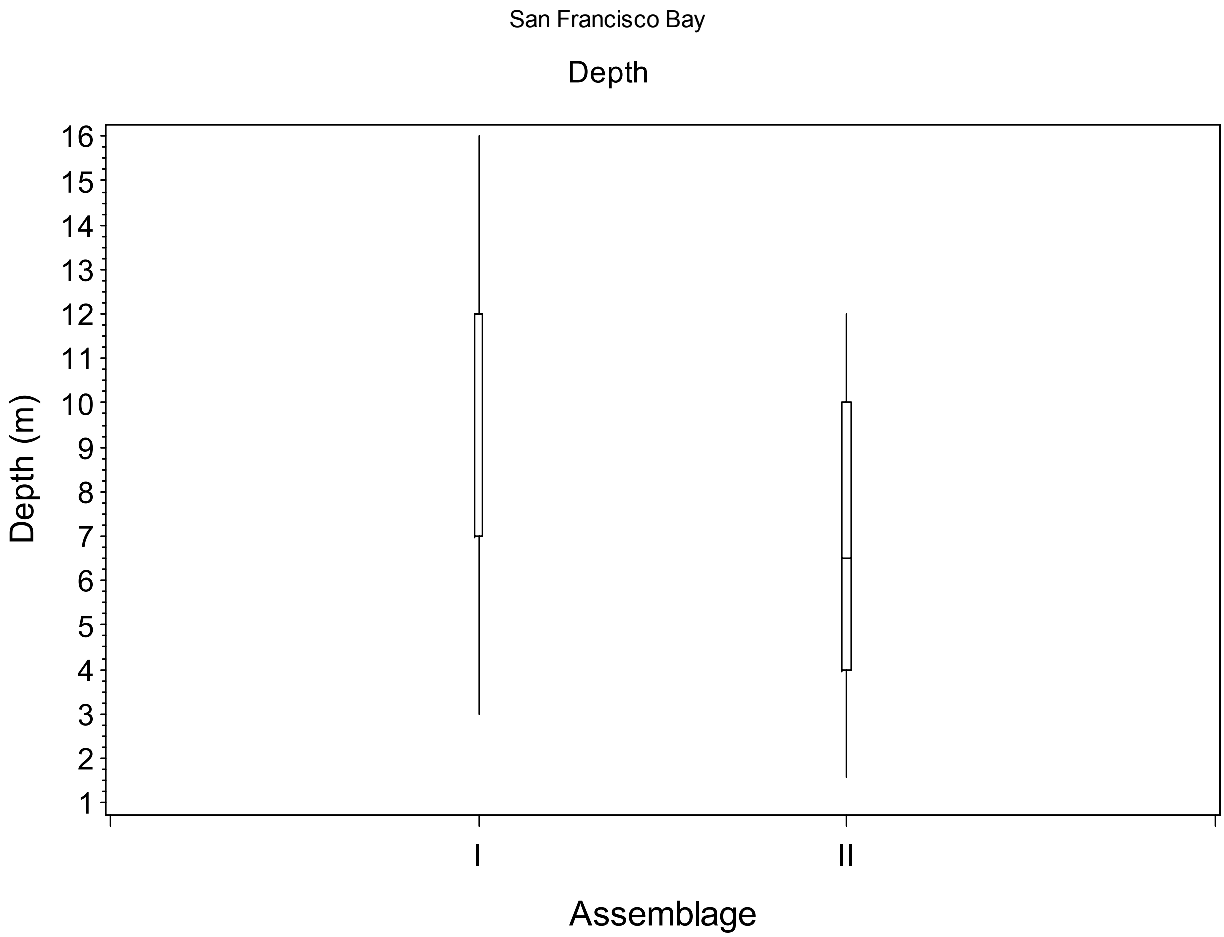

Figure 8. Box plot of depth distribution for each assemblage (sample group). See Figure 4 for an explanation.

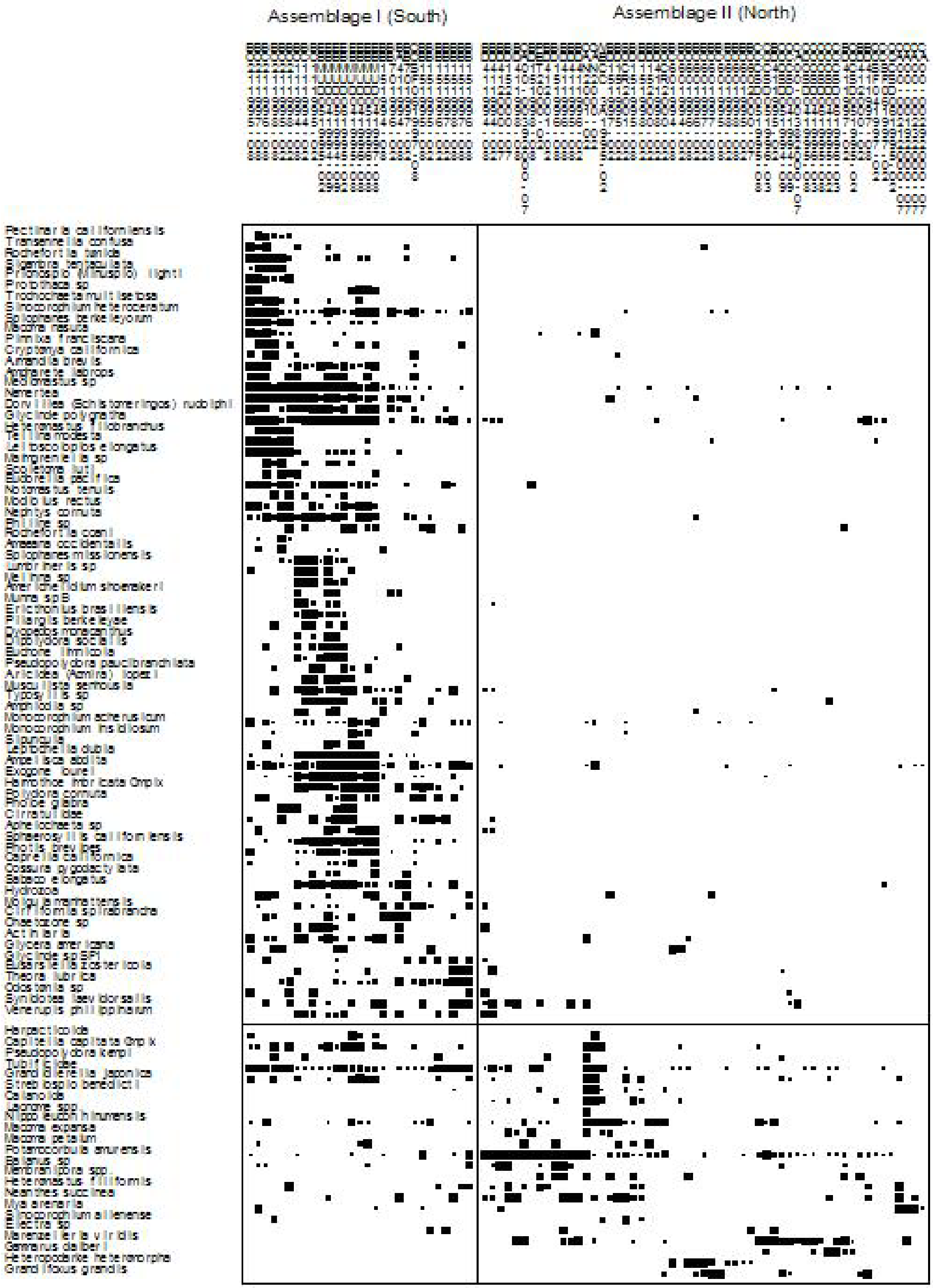

Figure 9. Two-way table showing species abundance differences among assemblages (sample groups).

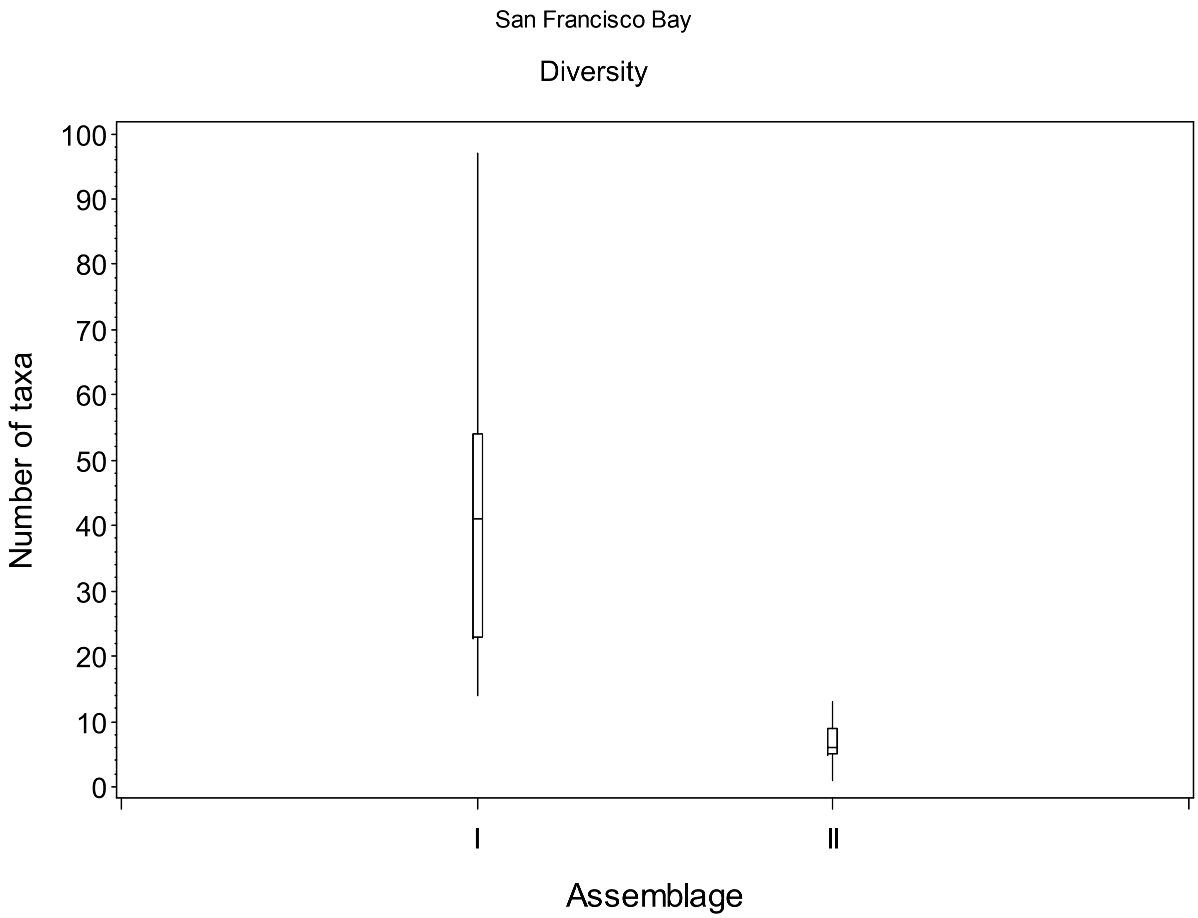

Figure 10. Box plot of diversity distribution for each assemblage (sample group). See Figure 4 for an explanation.

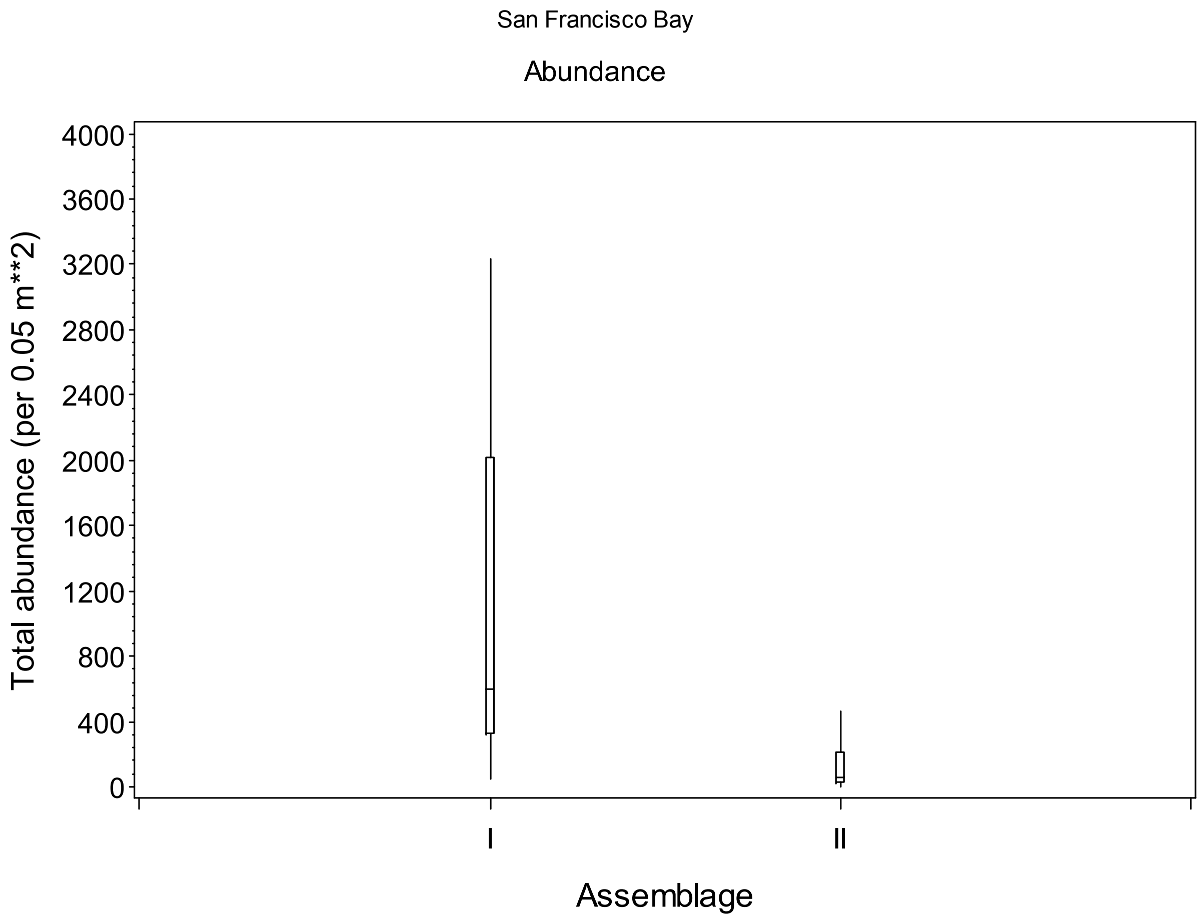

Figure 11. Box plot of abundance distribution for each assemblage (sample group). See Figure 4 for an explanation.

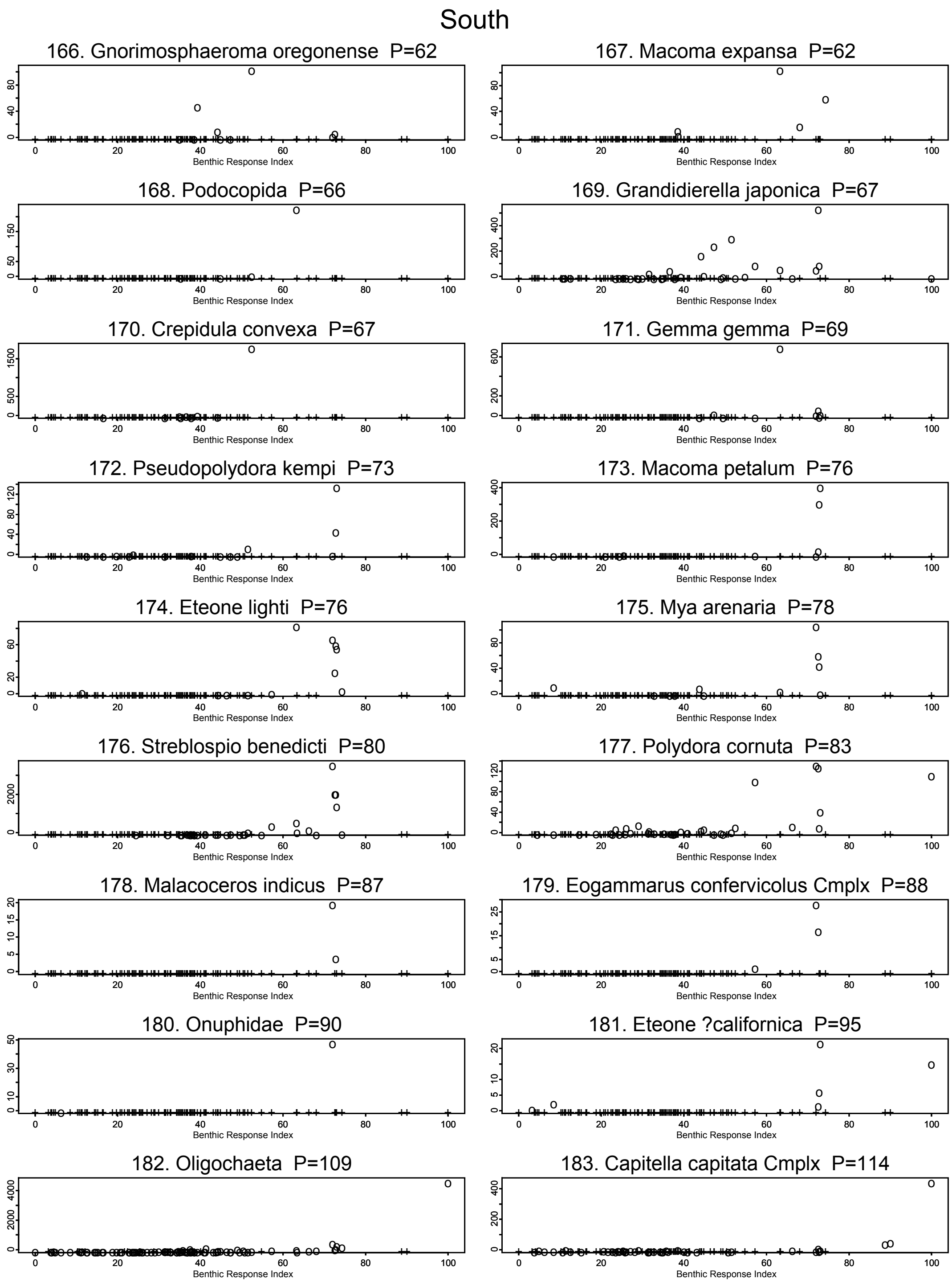

Figure 12. Abundances of species with the 18 highest pollution tolerance (p_i) scores in the southern San Francisco Bay habitat. Absences are indicated by plus symbols and occurrences by circles. Numbers to the left of species names are ranks of the pollution tolerance scores.

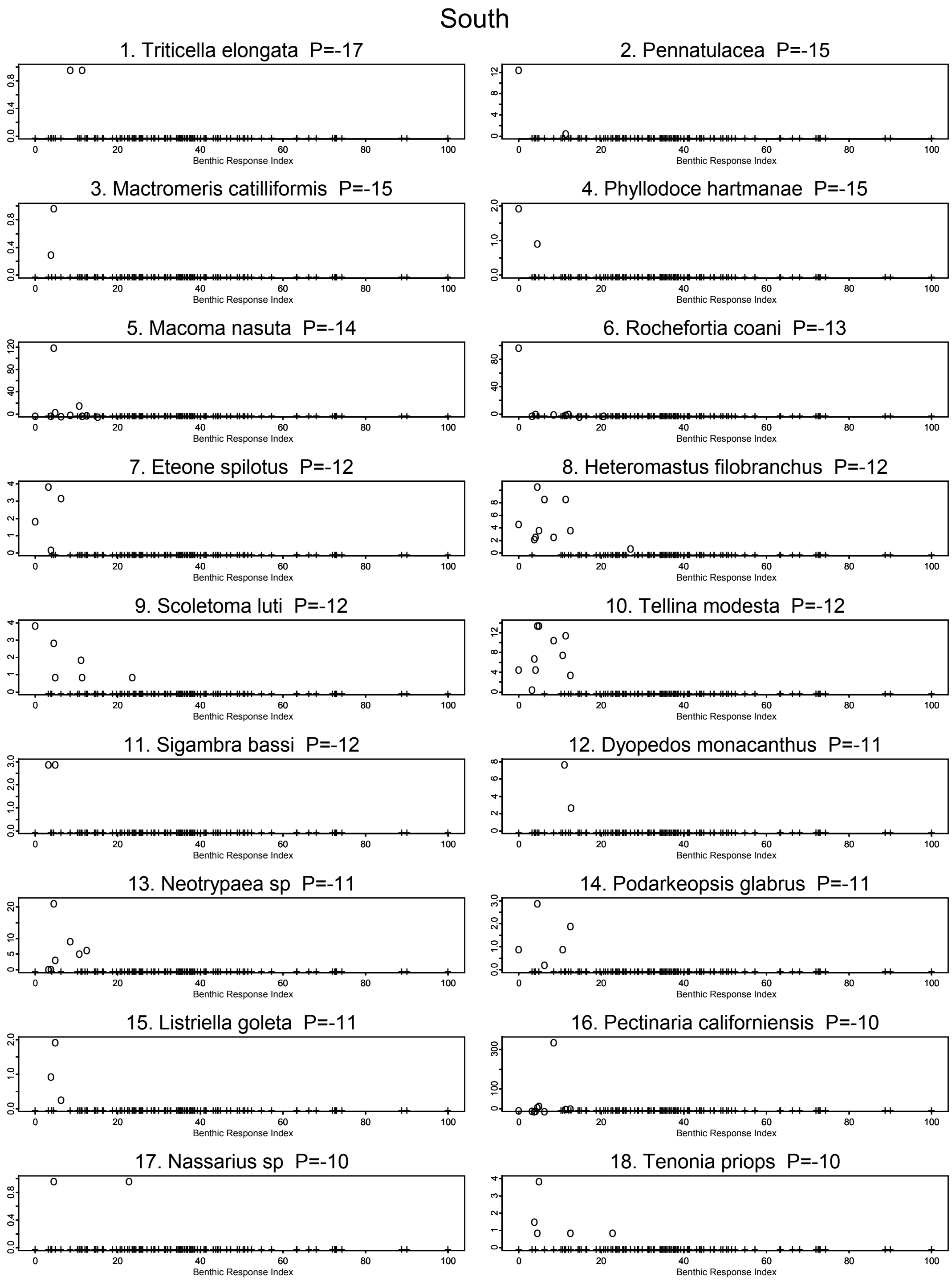

Figure 13. Abundances of species with the 18 lowest pollution tolerance (p_i) scores in the southern San Francisco Bay habitat. Absences are indicated by plus symbols and occurrences by circles. Numbers to the left of species names are ranks of the pollution tolerance scores.

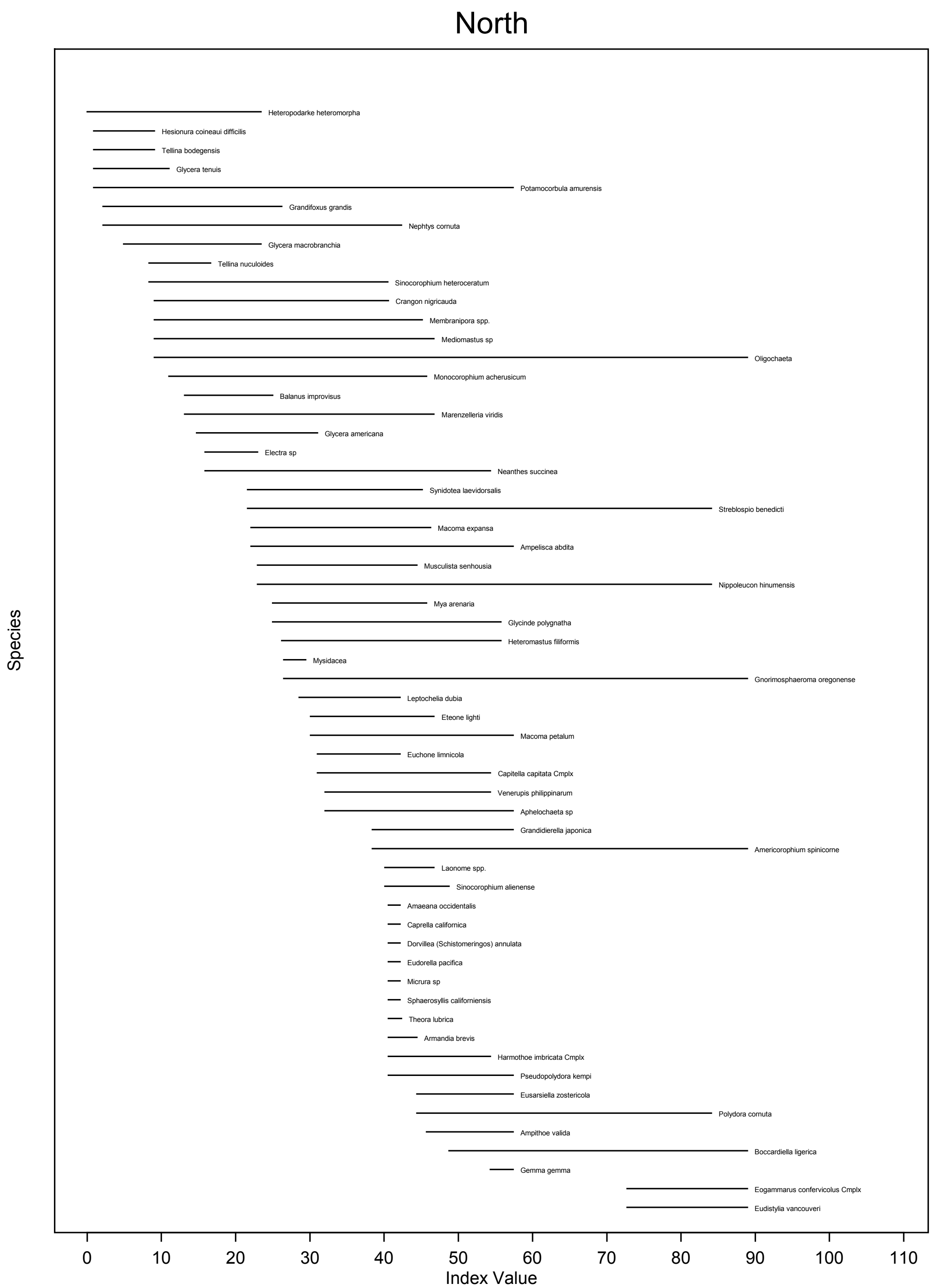

Figure 14. Species ranges along the index pollution gradient for the northern San Francisco Bay habitat. Species are ordered from top to bottom by their first and last appearance on the gradient. Only species occurring in at least two samples in the northern San Francisco Bay data are included.

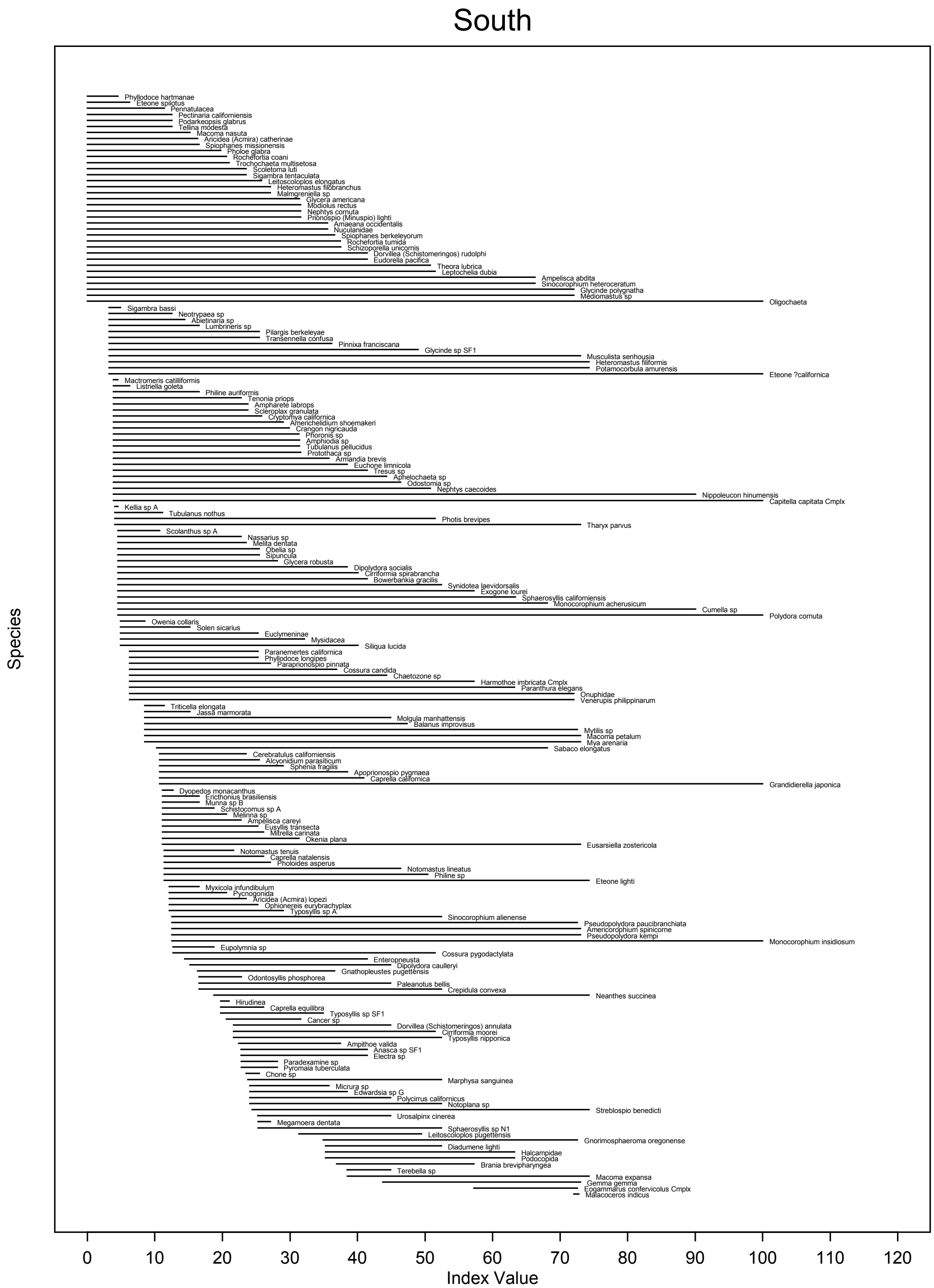

Figure 15. Species ranges along the index pollution gradient for the southern San Francisco Bay habitat. Species are ordered from top to bottom by their first and last appearance on the gradient. Only species occurring in at least two samples in the southern San Francisco Bay data are included.

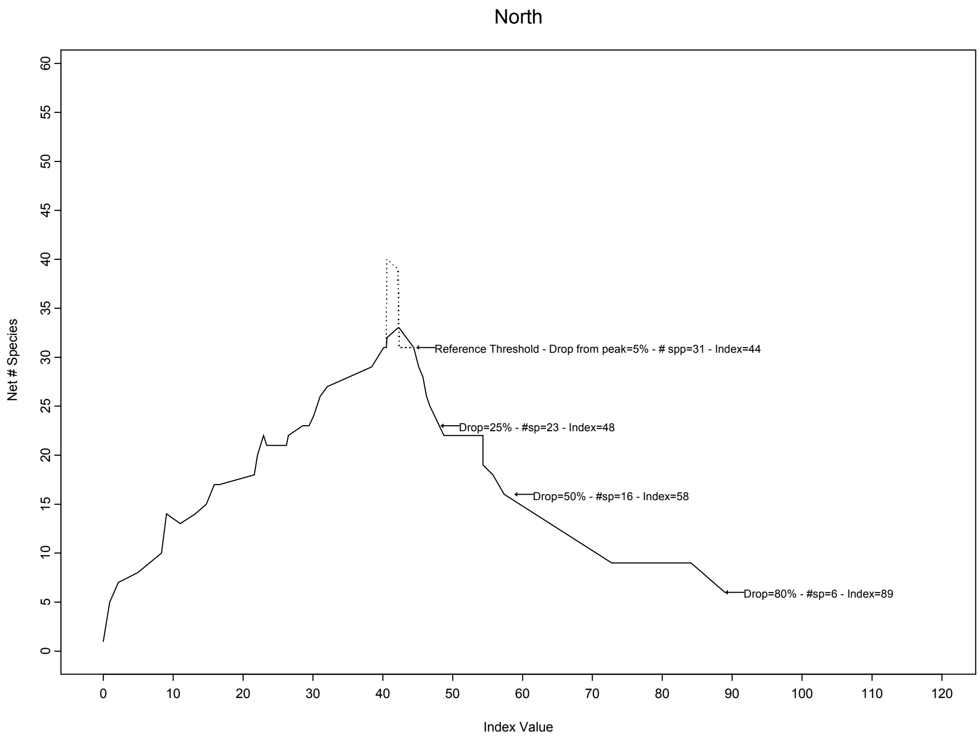

Figure 16. Summary of species ranges along the index pollution gradient for the northern San Francisco Bay habitat. The curve is the net number of species (cumulative number of species ranges intersecting index values up to and including the index value on the horizontal axis minus the cumulative number of species that have dropped out before the index value on the horizontal axis (see Figure 14). The outlier values indicated by dotted lines were eliminated before the peak number of species was determined. The labeled arrows indicate positions of the assessment thresholds on the curve.

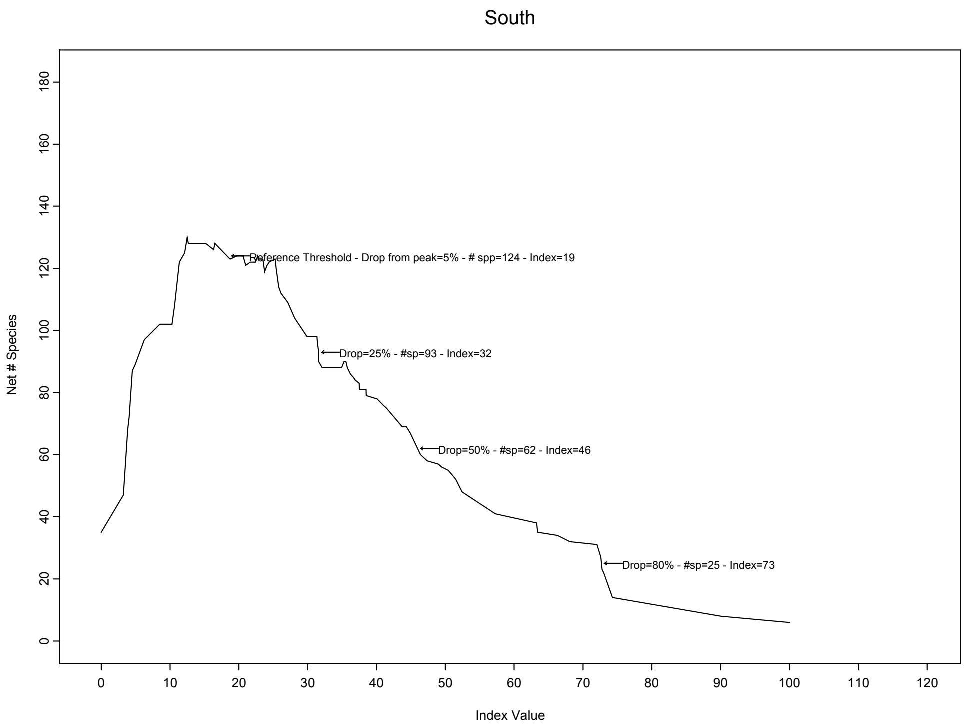

Figure 17. Summary of species ranges along the index pollution gradient for the southern San Francisco Bay habitat. The curve is the net number of species (cumulative number of species ranges intersecting index values up to and including the index value on the horizontal axis minus the cumulative number of species that have dropped out before the index value on the horizontal axis (see Figure 15). The labeled arrows indicate positions of the assessment thresholds on the curve.

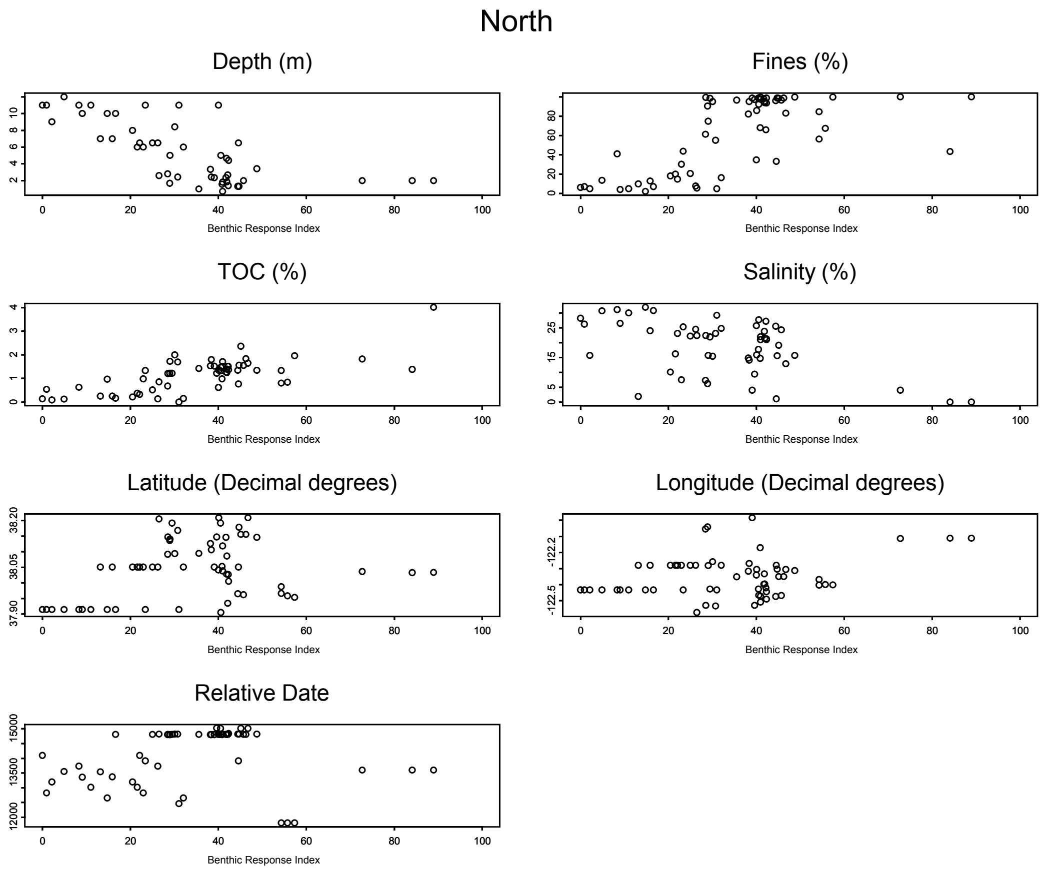

Figure 18. Relationships between habitat measures and index values for the northern San Francisco Bay habitat.

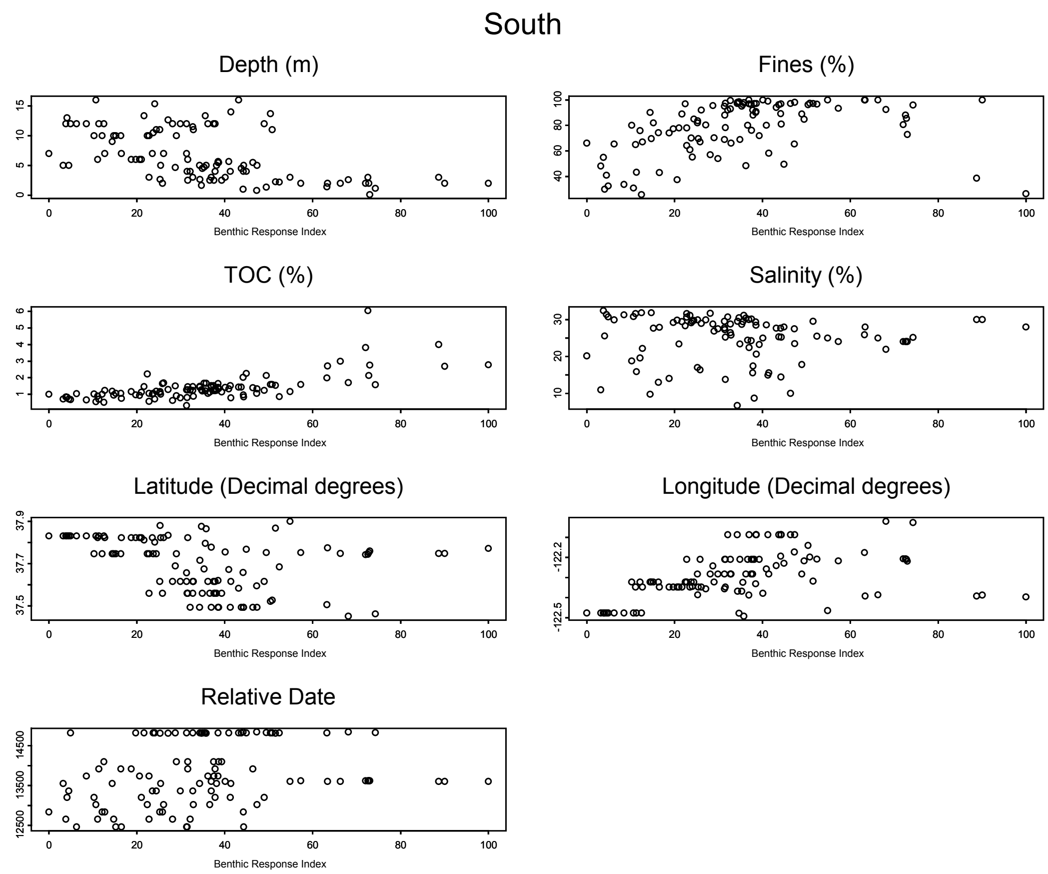

Figure 19. Relationships between habitat measures and index values for the southern San Francisco Bay habitat.

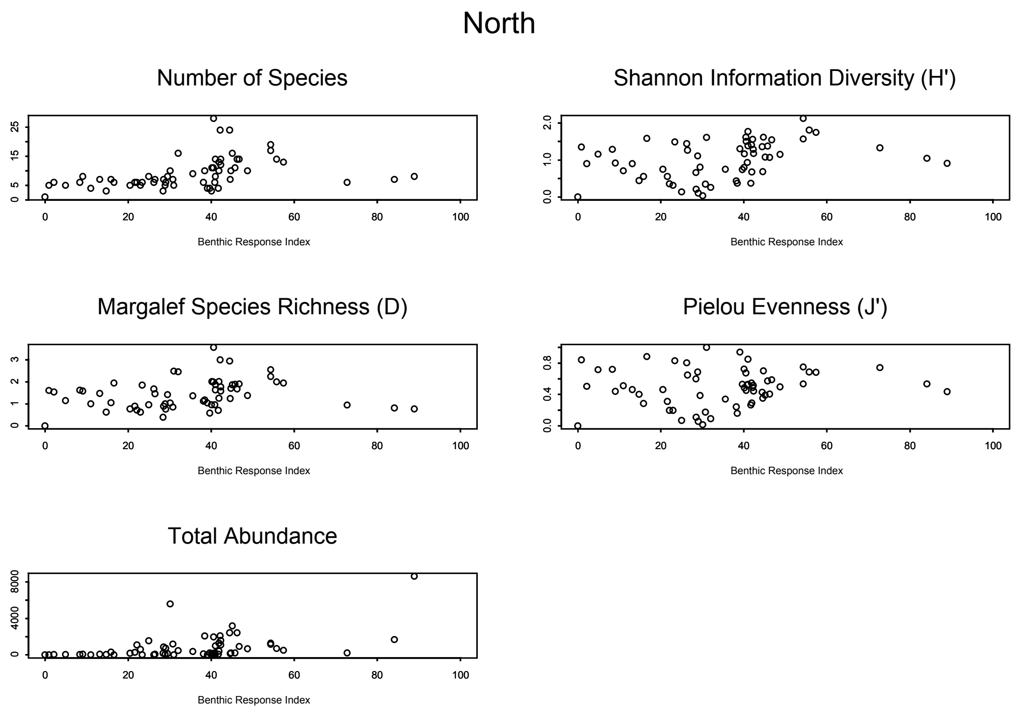

Figure 20. Relationships between biological community measures and index values for the northern San Francisco Bay habitat.

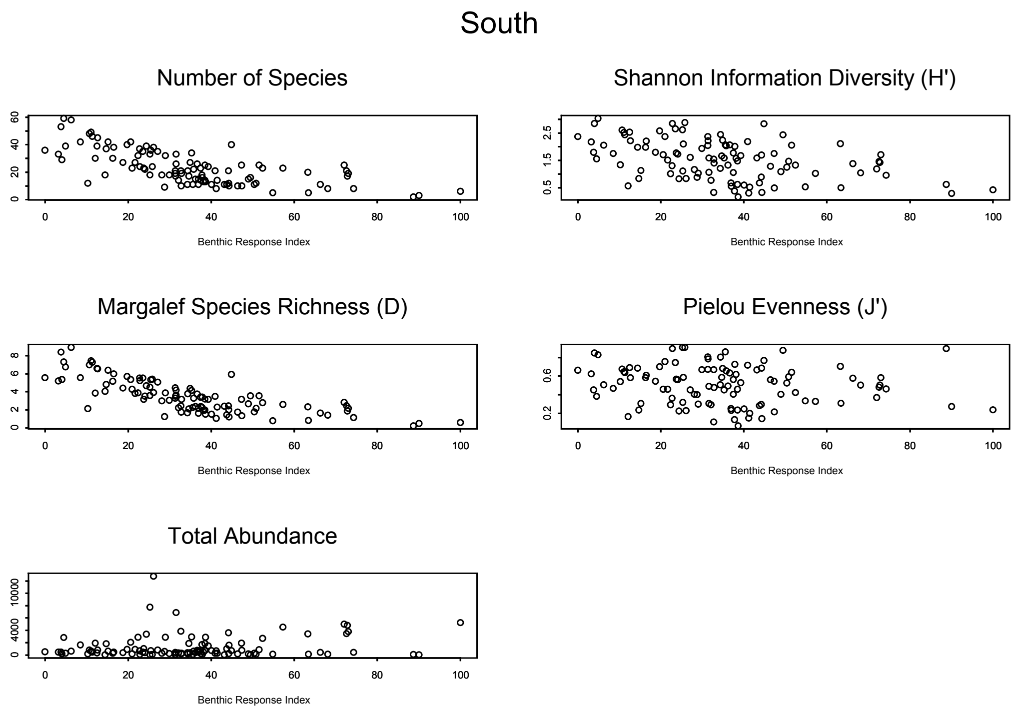

Figure 21. Relationships between biological community measures and index values for the southern San Francisco Bay habitat.

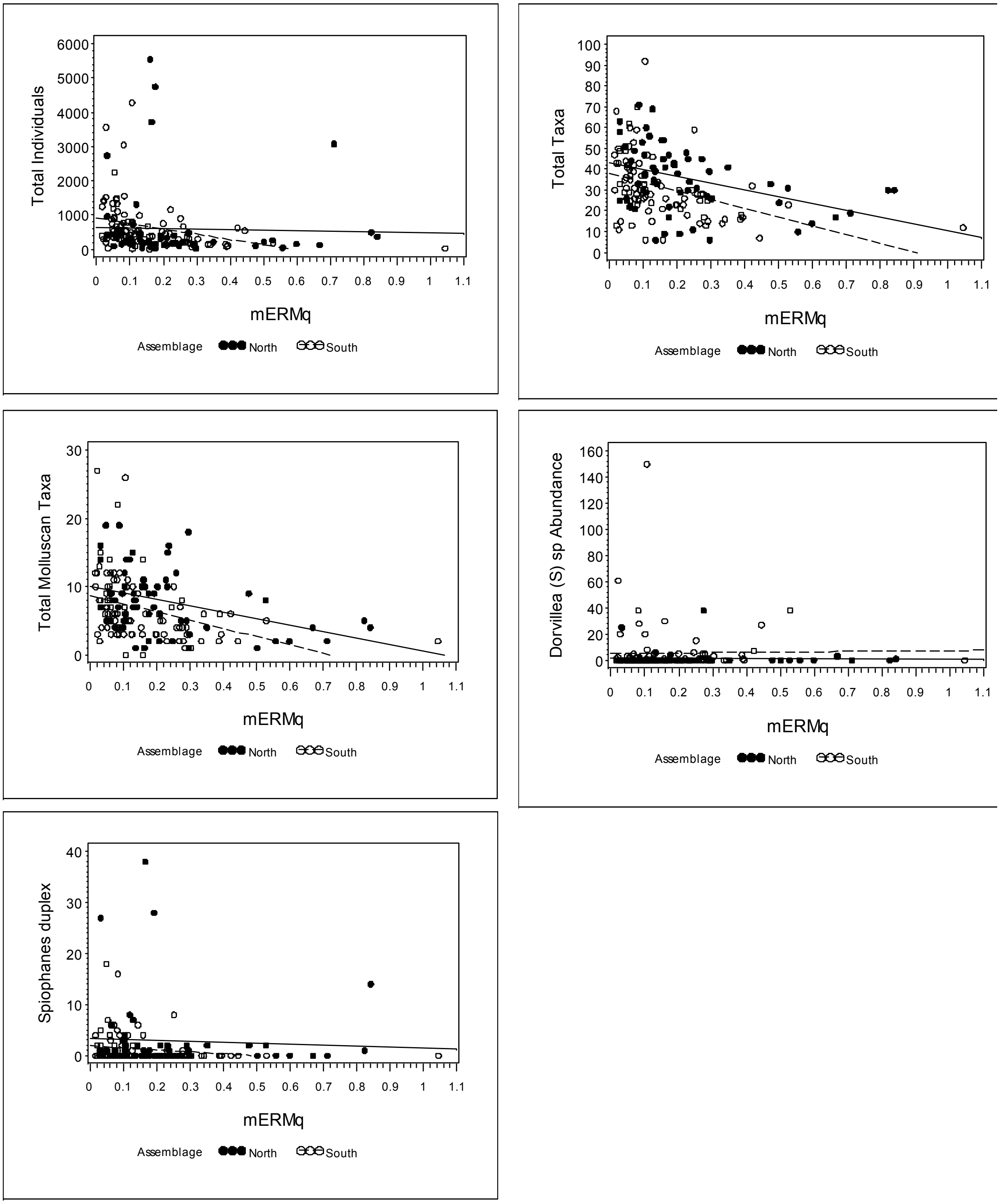

Figure 22. Relationships between selected assessment indicators and mean ERM quotients (mERMq) for the northern and southern bay habitats in southern California.

TABLES

Table 1. Benthic invertebrate samples in San Francisco Bay. Numbers of data and sampling details are presented for each source in the database. The sieve size for all samples was 0.5mm except DWR (0.595mm). C: Used in assessment comparison; D: Used in BRI development; A: All data; Reps: No. of pseudoreplicates; BADA: Bay Area Dischargers Association, see Bay Area Dischargers Association (1994); BPTCP: Bay Protection and Toxic Cleanup Program (see Hunt *et al.* 2001); CISNET: Coastal Intensive Sites Network (Thompson *et al.* 2002); DWR: Department of Water Resources (Department of Water Resources 1997); RMP: Regional Monitoring Program (Thompson *et al.* 1999); EMAP-ML: Environmental Monitoring and Assessment Program, collected by Moss Landing Marine Laboratories (Heitmuller *et al.* 1999); EMAP-NO: Environmental Monitoring and Assessment Program collected by National Oceanographic and Atmospheric Administration (Heitmuller *et al.* 1999).

Source	Frequency	Period	Area (m^2)	Reps	Samples			Sites		
					C	D	A	C	D	A
BADA	Wet & Dry Seasons	1994-1997	0.05	'94: 3			54			9
BPTCP-92	Once	May 1992	0.054 (3x0.018)		4	4	4	4	4	4
BPTCP-94	Once	Sep 1994	0.05				3			3
BPTCP-97	Once	Apr or Dec 1997	0.0225 (3x0.0075)		12	18	21	12	18	21
CISNET	Wet & Dry Seasons	2000-2001	0.05		8	8	12	4	4	4
DWR	Monthly	Jan 1994 – Dec 2001	0.053	4			904			15
RMP	Wet & Dry Seasons	1994-2000	0.05		68	87	101	7	8	9
RMP-W	Once	Feb-Mar 1995	0.05				4			4
EMAP-ML	Once	Jul-Aug 2000	0.05		17	17	17	17	17	17
EMAP-NO	Once	Jul-Aug 2000	0.044		33	33	33	33	33	33
			Total		**142**	**167**	**1153**	**77**	**84**	**119**

Table 2. Numbers of benthic data with synoptic sediment contaminant concentration and amphipod toxicity measurements. Chem: No. of sediment contaminant data; Tox: No. of amphipod toxicity data. See Table 1 for details about the data sources.

Source	No. of Samples			No. of Sites		
	All	Chem	Tox	All	Chem	Tox
BADA	54	53	0	9	9	0
BPTCP-92	4	4	4	4	4	4
BPTCP-94	3	0	0	3	0	0
BPTCP-97	21	18	21	21	18	21
CISNET	12	12	8	4	4	4
DWR	904	35	27	15	3	2
RMP	101	101	87	9	9	8
RMP-W	4	4	0	4	4	0
EMAP-ML	17	17	17	17	17	17
EMAP-NO	33	33	33	33	33	33
Total	1153	277	197	119	101	89

Table 3. Habitat criteria for each dendrogram split and the number and proportion of samples that met these criteria.

Split	Site Groups	N	Criterion	Met Criterion	
				N	Percent
1	I	29	North (Latitude > 37.9)	55	97.7
	II	57	South (Latitude <= 37.9)	29	

Table 4. Correlations between the ordination axes and the pollution indicator variables, after the canonical correlation analysis.

Analysis	Mean ERM Quotient		Amphipod Mortality	
	Axis 1	Axis 2	Axis 1	Axis 2
North	0.52	0.36	0.62	-0.26
South	0.78	0.00	0.22	0.48

Table 5. Optimum parameter values and index-pollution vector correlation coefficients. *f* is the exponent in the index calculations while *t* and *e* are only used to develop species pollution tolerance (p_i) values. *t* is the number of sites with only the *t* highest species abundance values included during index development. *e* is the exponent in the p_i calculations. $r_{I_s \cdot g_s}$ is the Pearson correlation between the optimized index and the pollution vector in the ordination space.

Habitat	*T*	*E*	*f*	$r_{I_s \cdot g_s}$
North	16	0.25	0.33	0.88
South	61	1.00	0.33	0.87

Table 6. Species with the 10 highest pollution tolerance scores in northern and southern San Francisco Bay. Included are species ranked in the top 10 in either habitat. The mean rank is the rank for the average of the pollution tolerance scores for northern and southern San Francisco Bay.

Phylum	Species	Pollution Tolerance Score		Rank		
		North	South	North	South	Mean
Annelida	*Eudistylia vancouveri*	109.43		2		1.0
Annelida	*Eteone ?californica*		94.59		3	1.5
Annelida	*Boccardiella ligerica*	90.96		4		2.0
Annelida	Onuphidae		89.74		4	2.0
Arthropoda	*Eogammarus confervicolus* Cmplx	111.34	88.45	1	5	3.0
Annelida	*Malacoceros indicus*		87.30		6	3.0
Annelida	Oligochaeta	62.36	108.62	10	2	6.0
Annelida	*Polydora cornuta*	71.77	83.26	7	7	7.0
Mollusca	*Gemma gemma*	65.31	68.71	9	13	11.0
Arthropoda	*Gnorimosphaeroma oregonense*	78.05	61.94	5	18	11.5
Annelida	*Streblospio benedicti*	47.19	80.23	18	8	13.0
Arthropoda	Americorophium spinicorne	93.24	56.31	3	24	13.5
Annelida	*Capitella capitata* Cmplx	38.83	114.23	28	1	14.5
Annelida	*Eteone lighti*	44.98	76.10	19	10	14.5
Mollusca	*Mya arenaria*	36.04	77.96	34	9	21.5
Arthropoda	*Eusarsiella zostericola*	68.80	31.28	8	44	26.0
Arthropoda	*Ampithoe valida*	72.05	20.16	6	70	38.0

Table 7. Species with the 10 lowest pollution tolerance scores in northern and southern San Francisco Bay. Included are species ranked in the top 10 in either habitat. The mean rank is the rank for the average of the pollution tolerance scores for northern and southern San Francisco Bay.

Phylum	Species	Pollution Tolerance Score		Rank		
		North	South	North	South	Mean
Annelida	*Hesionura coineaui difficilis*	-9.27		1		0.5
Ectoprocta	*Triticella elongata*		-17.35		1	0.5
Annelida	*Glycera macrobranchia*	-8.79		2		1.0
Cnidaria	Pennatulacea		-15.24		2	1.0
Annelida	*Glycera tenuis*	-3.96		3		1.5
Mollusca	*Mactromeris catilliformis*		-15.02		3	1.5
Annelida	*Phyllodoce hartmanae*		-14.89		4	2.0
Annelida	*Heteropodarke heteromorpha*	0.00		5		2.5
Mollusca	*Macoma nasuta*		-13.52		5	2.5
Mollusca	*Rochefortia coani*		-12.58		6	3.0
Mollusca	*Tellina bodegensis*	0.77		6		3.0
Annelida	*Eteone spilotus*		-12.48		7	3.5
Arthropoda	*Grandifoxus grandis*	7.29		7		3.5
Annelida	*Heteromastus filobranchus*		-11.94		8	4.0
Annelida	*Scoletoma luti*		-11.90		9	4.5
Ectoprocta	*Membranipora* spp.	10.78		10		5.0
Mollusca	*Tellina modesta*		-11.62		10	5.0
Annelida	*Glycera americana*	9.45	-5.14	9	44	26.5
Ectoprocta	*Electra* sp	9.12	20.93	8	116	62.0
Arthropoda	*Balanus improvisus*	-0.39	25.17	4	129	66.5

Table 8. Index threshold values for the northern and southern San Francisco Bay habitats.

Threshold	Northern Index Value	Southern Index Value
Reference	44	19
25% Biodiversity Loss	23	32
50% Biodiversity Loss	58	46
80% Biodiversity Loss	89	73

Table 9. Assessment of 142 samples as clearly disturbed or undisturbed by the San Francisco Benthic Response Index (SFBRI) developed for this study and the existing San Francisco Bay Index of Biotic Integrity (SFBIBI). Assessments by the two indices are significantly associated (p < 0.001; Mantel-Haenszel chi-square test).

		SFBRI	
		Disturbed	Undisturbed
SFBIBI	Disturbed	17	5
	Undisturbed	41	79

Table 10. Classification of 142 samples into categories of disturbance by the San Francisco Benthic Response Index (SFBRI) and the San Francisco Bay Index of Biotic Integrity (SFBIBI). Classifications by the two indices are significantly associated (p < 0.0001; Mantel-Haenszel chi-square test).

			SFBRI Classification				
			Reference	Response Level			
				1	2	3	4
SFBIBI Classification	Number of reference range exceedances	0	37	17	24	8	1
		1	14	11	6	2	0
		2	2	1	5	2	0
		3	0	1	1	1	1
		4	0	1	0	2	0
		5	0	0	0	4	1

Table 11. Reference samples identified for the North and South assemblages. *= Amphipod taxa without *Grandidierella japonica*, known to be tolerant.

Site	Location	Amphipod Toxicity	Tolerant/ Sensitive	Amphipod Taxa*	*Capitella capitata*	mERMq
			North			
B2152	Long Beach Harbor	0	0.46	7	0	0.0480
B2153	Long Beach Harbor	0	0.38	7	0	0.0324
B2154	Long Beach Harbor	0	0.47	4	0	0.0519
B2155	Long Beach Harbor	0	0.48	1	0	0.2376
B2156	Long Beach Harbor	0	0.44	1	0	0.2959
B2158	Los Angeles Harbor	0	0.50	3	0	0.1062
B2159	Los Angeles Harbor	0	0.43	4	0	0.1555
B2160	Los Angeles Harbor	0	0.50	1	0	0.2282
B2161	Los Angeles Harbor	0	0.44	2	0	0.2336
B2164	Anaheim Bay	0	0.38	3	0	0.1304
B2167	Long Beach Harbor	0	0.50	2	0	0.2591
B2175	Los Angeles Harbor	0	0.39	4	0	0.1616
B2184	Los Angeles Harbor	0	0.39	2	0	0.4773
B2186	Los Angeles Harbor	0	0.43	7	0	0.1288
B2187	Long Beach Harbor	0	0.43	6	0	0.0881
B2426	Los Angeles Harbor	0	0.45	4	0	0.0649
E3031	King Harbor	0	0.40	9	0	0.0323
E3034	Long Beach Harbor	0	0.47	4	0	0.0751
E3037	Long Beach Harbor	0	0.46	4	0	0.1105
			South			
B2227	San Diego Bay	0	0.41	6	0	0.0532
B2228	San Diego Bay	0	0.42	5	0	0.0966
B2229	San Diego Bay	0	0.38	6	0	0.0723
B2231	San Diego Bay	0	0.38	9	0	0.0598
B2233	San Diego Bay	0	0.45	3	0	0.0597
B2252	San Diego Bay	0	0.35	6	0	0.0316
B2263	San Diego Bay	0	0.45	5	0	0.1438
B2265	San Diego Bay	0	0.37	6	0	0.0284
B2423	Mission Bay	0	0.47	4	0	0.0220
B2434	San Diego Bay	0	0.43	3	0	0.0723
B2435	San Diego Bay	0	0.35	3	0	0.0321
B2436	San Diego Bay	0	0.40	8	0	0.0893
B2441	San Diego Bay	0	0.40	4	0	0.0837
B2442	San Diego Bay	0	0.43	3	0	0.1594
SDR03	San Diego Bay	0	0.50	2	0	0.0738

Table 12A. Results of IBI assessment of southern California samples from the North assemblage. Hit = 1 indicates that the reference range was exceeded, Hit = 0 indicates that the reference range was not exceeded.

North	Location	Total Taxa	Hit	Molluscan Taxa	Hit	*Dorvillea (S.) sp.*	Hit	*Spiophanes duplex*	Hit	Assessment Value	mERMq
B2128	Pierpont Bay	10	1	2	1	0	0	0	0	2	0.5585
B2129	Channel Islands Harbor	30	0	5	0	0	0	1	0	0	0.8241
B2130	Channel Islands Harbor	24	1	1	1	0	0	0	0	2	0.5026
B2131	Channel Islands Harbor	22	1	5	0	0	0	0	0	1	0.0614
B2134	Anaheim Bay	29	0	6	0	0	0	2	0	0	0.2116
B2157	Long Beach Harbor	47	0	12	0	0	0	3	0	0	0.104
B2163	Long Beach Harbor	38	0	10	0	0	0	0	0	0	0.2028
B2169	Los Angeles Harbor	14	1	2	1	0	0	0	0	2	0.5997
B2170	Los Angeles Harbor	30	0	4	1	1	0	14	0	1	0.8424
B2172	Los Angeles Harbor	43	0	7	0	0	0	28	1	1	0.1923
B2173	Los Angeles Harbor	30	0	10	0	0	0	2	0	0	0.2327
B2174	Los Angeles Harbor	41	0	10	0	0	0	38	1	1	0.1659
B2176	Los Angeles Harbor	21	1	4	1	0	0	0	0	2	0.0757
B2178	Los Angeles Harbor	31	0	8	0	0	0	2	0	0	0.5281
B2179	Los Angeles Harbor	30	0	4	1	0	0	0	0	1	0.1011
B2185	Los Angeles Harbor	39	0	8	0	0	0	0	0	0	0.1365
B2188	Anaheim Bay	34	0	8	0	0	0	0	0	0	0.1335
B2421	Los Angeles Harbor	45	0	5	0	38	1	0	0	1	0.2744
B2427	Los Angeles Harbor	27	0	5	0	0	0	2	0	0	0.2908
B2430	Los Angeles Harbor	41	0	4	1	0	0	2	0	1	0.3513
B2431	Los Angeles Harbor	33	0	7	0	0	0	2	0	0	0.1419
B2432	Los Angeles Harbor	17	1	2	1	0	0	0	0	2	0.1768
B2443	Marina Del Rey	6	1	1	1	0	0	0	0	2	0.297
B2444	Marina Del Rey	6	1	1	1	0	0	0	0	2	0.1375
B2445	Marina Del Rey	11	1	2	1	0	0	0	0	2	0.248
B2446	Marina Del Rey	9	1	1	1	0	0	0	0	2	0.1637
B2447	Marina Del Rey	22	1	6	0	0	0	1	0	1	0.1775
B2448	Marina Del Rey	45	0	11	0	0	0	0	0	0	0.1609
B2449	Marina Del Rey	35	0	9	0	6	0	0	0	0	0.1323
B2450	Long Beach Harbor	9	1	2	1	0	0	0	0	2	0.2085
B2451	Long Beach Harbor	47	0	10	0	0	0	4	0	0	0.1045
E3026	Santa Barbara	47	0	9	0	4	0	0	0	0	0.1762
E3028	Channel Islands Harbor	19	1	2	1	0	0	0	0	2	0.7128
E3029	Channel Islands Harbor	33	0	7	0	2	0	0	0	0	0.0852
E3030	Point Mugu	25	1	7	0	0	0	1	0	1	0.032
E3032	Los Angeles Harbor	17	1	4	1	3	0	0	0	2	0.6696
E3033	Los Angeles Harbor	56	0	14	0	0	0	8	0	0	0.1195
E3035	Long Beach Harbor	42	0	9	0	0	0	0	0	0	0.1929
E3036	Los Angeles Harbor	37	0	5	0	1	0	1	0	0	0.105
E3038	Los Angeles Harbor	53	0	9	0	0	0	3	0	0	0.0982
E3039	Los Angeles Harbor	26	0	3	1	0	0	0	0	1	0.2993

Table 12B. Results of IBI assessment of southern California samples from the South assemblage. Hit = 1 indicates that the reference range was exceeded, Hit = 0 indicates that the reference range. Hit = 1 indicates that the reference range was exceeded, Hit = 0 indicates that the reference range was not exceeded.

South	Location	Total Taxa	Hit	Total Abundance	Hit	Molluscan Taxa	Hit	Assessment Value	mERMq
B2221	San Diego Bay	28	1	781	0	6	1	2	0.082
B2222	San Diego Bay	32	0	684	0	4	1	1	0.1539
B2223	San Diego Bay	30	0	810	0	5	1	1	0.1081
B2224	San Diego Bay	37	0	380	0	9	0	0	0.0627
B2225	San Diego Bay	59	0	3049	1	11	0	1	0.0815
B2226	San Diego Bay	46	0	993	0	10	0	0	0.1291
B2230	San Diego Bay	60	0	1345	0	12	0	0	0.0622
B2235	San Diego Bay	25	1	539	0	4	1	2	0.0523
B2238	San Diego Bay	35	0	746	0	7	1	1	0.047
B2239	San Diego Bay	23	1	1016	0	4	1	2	0.0839
B2240	San Diego Bay	36	0	1186	0	12	0	0	0.0508
B2241	San Diego Bay	40	0	1499	0	8	0	0	0.0604
B2242	San Diego Bay	25	1	1112	0	3	1	2	0.0635
B2243	San Diego Bay	41	0	947	0	8	0	0	0.0537
B2244	San Diego Bay	44	0	1339	0	9	0	0	0.0489
B2245	San Diego Bay	22	1	485	0	5	1	2	0.0631
B2247	San Diego Bay	29	0	890	0	6	1	1	0.0453
B2249	San Diego Bay	35	0	590	0	7	1	1	0.063
B2251	San Diego Bay	26	1	1166	0	3	1	2	0.2226
B2253	San Diego Bay	30	0	452	0	6	1	1	0.2072
B2254	San Diego Bay	28	1	674	0	4	1	2	0.2606
B2255	San Diego Bay	26	1	385	0	6	1	2	0.1675
B2256	San Diego Bay	26	1	234	1	5	1	3	0.1026
B2257	San Diego Bay	33	0	495	0	7	1	1	0.1046
B2258	San Diego Bay	30	0	806	0	4	1	1	0.0938
B2259	San Diego Bay	20	1	96	1	5	1	3	0.1168
B2260	San Diego Bay	42	0	2249	1	6	1	2	0.0549
B2262	San Diego Bay	25	1	529	0	4	1	2	0.0867
B2264	San Diego Bay	25	1	233	1	4	1	3	0.2688
B2424	Mission Bay	47	0	396	0	12	0	0	0.0164
B2425	Mission Bay	92	1	4286	1	26	0	2	0.1067
B2433	San Diego Bay	51	0	681	0	10	0	0	0.0593
B2438	San Diego Bay	31	0	358	0	7	1	1	0.0515
B2439	San Diego Bay	28	1	527	0	5	1	2	0.1225
B2440	San Diego Bay	49	0	596	0	5	1	1	0.0493
E3041	Santa Margarita River	30	0	1252	0	10	0	0	0.0172
E3042	Agua Hedionda Lagoon	50	0	213	1	8	0	1	0.0279
E3043	Agua Hedionda Lagoon	43	0	304	0	12	0	0	0.0215
E3044	Mission Bay	13	1	299	0	6	1	2	0.0675
E3045	San Diego River	11	1	3568	1	2	1	3	0.0292

Table 12B (continued).

B. South	Location	Total Taxa	Hit	Total Abundance	Hit	Molluscan Taxa	Hit	Assessment Value	mERMq
E3046	San Diego River	13	1	1411	0	3	1	2	0.0228
E3047	San Diego Bay	26	1	364	0	3	1	2	0.0874
E3048	San Diego Bay	15	1	579	0	3	1	2	0.1266
E3049	San Diego Bay	26	1	724	0	5	1	2	0.0752
E3050	San Diego Bay	23	1	523	0	3	1	2	0.197
SDC01	San Diego Bay	28	1	368	0	7	1	2	0.2716
SDC02	San Diego Bay	25	1	150	1	4	1	3	0.2901
SDC03	San Diego Bay	16	1	158	1	2	1	3	0.3877
SDC04	San Diego Bay	23	1	461	0	5	1	2	0.2195
SDC05	San Diego Bay	14	1	202	1	2	1	3	0.2683
SDC06	San Diego Bay	28	1	298	0	4	1	2	0.2782
SDC07	San Diego Bay	29	0	414	0	3	1	1	0.1018
SDC08	San Diego Bay	6	1	20	1	0	1	3	0.1079
SDC09	San Diego Bay	32	0	630	0	6	1	1	0.4229
SDC10	San Diego Bay	26	1	309	0	1	1	2	0.3042
SDC11	San Diego Bay	6	1	6	1	0	1	3	0.1593
SDC12	San Diego Bay	12	1	32	1	2	1	3	1.0457
SDC13	San Diego Bay	23	1	188	1	5	1	3	0.53
SDC14	San Diego Bay	7	1	549	0	2	1	2	0.4448
SDP01	San Diego Bay	29	0	152	1	10	0	1	0.11
SDP02	San Diego Bay	21	1	124	1	5	1	3	0.2108
SDP03	San Diego Bay	24	1	247	0	3	1	2	0.1249
SDP04	San Diego Bay	18	1	202	1	2	1	3	0.231
SDP05	San Diego Bay	15	1	126	1	3	1	3	0.2917
SDP06	San Diego Bay	13	1	69	1	1	1	3	0.286
SDP07	San Diego Bay	17	1	176	1	2	1	3	0.2809
SDP08	San Diego Bay	23	1	758	0	3	1	2	0.2002
SDP09	San Diego Bay	15	1	38	1	4	1	3	0.0348
SDP10	San Diego Bay	21	1	246	0	3	1	2	0.1577
SDP11	San Diego Bay	18	1	83	1	6	1	3	0.3904
SDP12	San Diego Bay	30	0	302	0	7	1	1	0.2455
SDP13	San Diego Bay	27	1	758	0	6	1	2	0.1076
SDP14	San Diego Bay	28	1	482	0	8	0	1	0.2771
SDP15	San Diego Bay	17	1	113	1	3	1	3	0.3945
SDP16	San Diego Bay	14	1	151	1	2	1	3	0.3346
SDP17	San Diego Bay	16	1	150	1	6	1	3	0.3423
SDR02	San Diego Bay	31	0	684	0	6	1	1	0.0753
SDR04	San Diego Bay	59	0	904	0	10	0	0	0.2521
SDR05	San Diego Bay	47	0	453	0	11	0	0	0.1139
SDR06	San Diego Bay	28	1	410	0	3	1	2	0.0817

Table 13A. Candidate indicators and indicator selection criteria for the North assemblage. N: Number of samples in which the indicator occurred; HiLo p identifies whether the indicator was in either the upper or lower 10 BRI pollution tolerance scores (see text); Correlated = Spearman's rank correlations (α < 0.05); ** means the indicator was used in San Francisco Estuary assessments; * means that mERMq was a significant contributor in the multiple regression analyses (see Table 2); + = selected for optimization testing.

Indicator	N	HiLo p	Correlation with mERMq	Proportion R^2 from mERMq
Total Taxa ** +	60		+	0.075
Total Abundance ** +	60		+	0.002
Amphipod taxa **	55		+	0.138
Molluscan taxa ** +	60		+	0.267
Capitellidae +	53		+	0.238*
Dorvillidae	8		+	0.247
Capitellla capitata complex **	6	L	-	0.170*
Spiophanes benedicti **	5		-	0.041
Grandidierella japonica	13		+	0.017
Mediomastus sp.	44		+	0.089
Amphipholis spp	4		-	0.639
Dorvillea (Schistomeringos) sp. +	8		-	0.250
Musculista senhousia	44	H	-	0.514
Eteone spp	3	L	-	0.044
Acteocina inculta	5	L	-	0.032
Ericthonius brasiliensis	5	L	-	0.018
Leitoscoloplos pugettensis	46		-	0.128
Goniada littorea	4		+	0.142
Notomastus sp. +	33		+	0.085
Amphideutopus oculatus	30		+	0.002
Spiophanes duplex +	23		-	0.211
Tagelus subteres	30		+	0.123

Table 13B. Candidate indicators and indicator selection criteria for the South assemblage. FO = frequency of occurrence, HiLo p identifies whether the indicator was in either the upper or lower 10 BRI pollution tolerance scores (see text), Correlated = Spearman's rank correlations (α < 0.05), ** means the indicator was used in San Francisco Estuary assessments, * means that mERMq was a significant contributor in the multiple regression analyses (see Table 2), + = selected for optimization testing.

Indicator	FO	HiLo p	Correlation with mERMq	Proportion R^2 from mERMq
Total Taxa * +	95		+	0.303
Total Abundance * +	95		+	0.027
Amphipod taxa *	84		+	0.004
Molluscan taxa * +	93		+	0.548*
Capitellidae	93		+	0.224
Dorvillidae +	35		+	0.208*
Capitellla capitata complex *	17		-	
Spiophanes benedicti *	5		-	
Grandidierella japonica	14		+	0.187
Mediomastus sp. +	89		+	0.597
Amphipholis spp	21		+	0.361
Dorvillea (Schistomeringos) sp.	35		+	0.133*
Musculista senhousia	82		+	0.084
Eteone spp	22		+	0.286
Acteocina inculta +	32		+	0.540
Ericthonius brasiliensis	10		+	0.033
Leitoscoloplos pugettensis	85		+	0.121
Goniada littorea	14		+	0.124
Notomastus sp. +	15		+	0.561*
Amphideutopus oculatus	63		+	0.276
Spiophanes duplex +	28		+	0.874
Tagelus subteres	37		+	0.581*

Table 14. Results of the multiple regression analysis of benthic assessment indicators for Southern California. Superscript a = arcsin transformation, L=log transformation, * means significant values (p<0.05).

	Partial Coefficients				
	Independent Variables				Proportion from mERMq
Indicator	**Fines**	**TOC**	**mERMq**	**Total R^2**	
North					
Total Taxa	0.216^{a*}	0.046^{a}	0.020^{L}	0.267*	0.075
Log Molluscan taxa	0.145^{a}	0.017^{a}	0.055^{L}	0.206*	0.267
Log *Dorvillea (S.) sp.*	0.026^{a}	0.079*	0.033^{L}	0.131*	0.250
Log *Spiophanes duplex*	0.213*	0.007^{a}	0.055^{L}	0.261*	0.211
South					
Total Taxa	0.005	0.112*	0.047^{L}	0.158*	0.303
Total Abundance	0.035*	0.346*	0.010^{L}	0.375*	0.027
Molluscan taxa	0.003	0.158*	0.163^{L*}	0.298*	0.548

Table 15. Reference ranges for the benthic indicators used in the southern California assessment.

	North		South		SF Bay Polyhaline	
Indicator	**minimum**	**maximum**	**minimum**	**maximum**	**minimum**	**maximum**
Total taxa	26	71	29	70	21	66
Total Abundance			241	1560	97	2931
Molluscan taxa	5	19	8	27		
Dorvillea (S.) sp	0	25				
Spiohhanes duplex	0	27				
MERMq	0.032	0.477	0.022	0.159	0.048	0.146

Table 16. Impacted samples identified for southern California using the San Francisco Estuary assessment method *per se*. Hit = 1 indicates that the reference range was exceeded, Hit = 0 indicates that the reference range was not exceeded. AV = assessment value.

Site	- Total Taxa	Hit	Total Abund.	Hit	Amphipod Taxa	Hit	*Capitella capitata*	Hit	AV (#)
					North				
B2423	17	0	29	1	0	1	0	0	2
B2443	6	1	35	1	0	1	0	0	3
B2444	6	1	42	0	0	1	0	0	2
B2446	9	1	76	0	0	1	0	0	2
B2450	9	1	160	0	1	0	101	1	2
					South				
SDC03	6	1	20	1	0	0	0	0	2
SDC11	6	1	6	1	0	0	0	0	2
SDC12	12	1	32	1	0	0	0	0	2
SDC14	7	1	549	0	1	0	501	1	2
E3045	11		3568	0	4	0	380	1	2

Table 17. Comparison of abiotic variables in reference and impacted samples, Wilcoxon 2-sample test, *= significantly different.

Variable	Reference mean	n	Impacted mean	n	p
		North			
Percent Fines	60.3	19	82.4	12	0.013*
TOC	1.22	19	2.16	12	0.040*
mERMq	0.154	19	0.363	12	0.008*
		South			
Percent Fines	46.3	15	53.3	49	0.210
TOC	0.97	14	1.18	22	0.361
mERMq	0.072	15	0.209	49	0.0003*

Table 18. Percentage of impacted samples in each of several mERMq ranges for IBI and BRI assessments.

	Percent Impacted	
mERMq range	**IBI**	**BRI**
	North	
<0.061	0.0	0.0
0.062 - 0.1375	9.1	9.1
0.137 - 0.352	35.0	45.0
>0.502	62.5	75.0
	South	
<0.0215	0.0	0.0
0.023 - 0.063	27.8	72.2
0.0631 - 0.114	61.9	90.5
0.116 - 0.423	80.0	94.3
>0.444	100.0	100.0

Table 19. Classification of 155 samples as impacted or unimpacted by the southern California IBI (SC IBI) used in this study and the southern California BRI (SC BRI). Assessments by the two indices are significantly associated (p < 0.001; Mantel-Haenszel chi-square test).

		SC BRI	
		Impacted	**Unimpacted**
SC IBI	**Impacted**	56	5
	UnImpacted	29	65

Table 20. Classification of 155 samples in categories of impact by the southern California IBI (SC IBI) used in this study and the southern California BRI (SC BRI). Assessments by the two indices are significantly associated (p < 0.0001; Mantel-Haenszel chi-square test).

			SC BRI Response Level				
			0	**1**	**2**	**3**	**4**
SC IBI	**Assessment Value**	**0**	57	8	1	0	0
		1	8	9	8	3	0
		2	4	16	12	5	3
		3	1	2	7	11	0
		4	0	0	0	0	0

APPENDIX A

METHOD FOR FINDING THE POLLUTION GRADIENT IN ORDINATION SPACE

An example of the method used to find the pollution gradient in the ordination space is presented here. Canonical correlation analysis was used to reduce the multivariate ordination space to a two-dimensional space that maximally correlates with the pollution gradient. The canonical correlation analysis used the first 20 ordination axes and the two pollution indicator variables, the mean ERM quotient and the control-adjusted amphipod mortality in acute sediment toxicity tests. The canonical correlation analysis produces two-dimensional spaces, one corresponding to the ordination scores and the other corresponding to the indicator variables. The space used for index development corresponds to the ordination scores.

Table A1 presents example correlations between the first and second canonical correlation axes and the pollution indicators. The correlations for the amphipod toxicity test are represented graphically on the left of Figure A1 as distances along Axis 1 and Axis 2. The resultant direction or vector for the amphipod test is 43.5^{o} from Axis 1, indicated by a line crossing through the origin and a bisection of the line connecting the two correlations. Using the same method, on the right of Figure A1, the resultant vector for the mean ERM quotient is found to be at -15^{o}. The overall pollution gradient vector is computed as the average of the two vectors for the pollution indicators, i.e., 14.25 from the horizontal ((43.5-15.0)/2).

Table A1. Example correlations between the indicators and the ordination axes after the canonical correlation analysis. These correlations are presented graphically in Figure A1.

	Axis 1	Axis 2
Mean ERM Quotient	0.82	-0.22
Amphipod Mortality	0.52	0.49

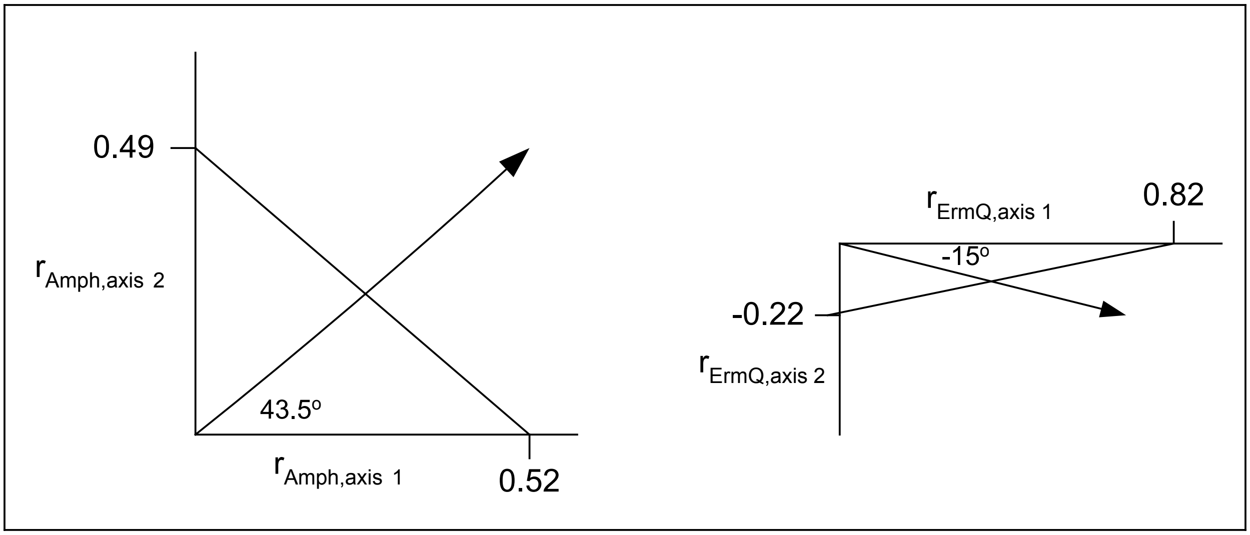

Figure A1. Example of the method for finding resultant vectors for pollution indicators in a two-dimensional ordination space, using the correlations in Table A.1. See text for explanation.