AQUATIC SCIENCE CENTER

www.aquaticsciencecenter.org

CONTRIBUTION NO. 584 • JUL 09

Estimated Sediment Contaminant Concentrations Associated with Biological Impacts at San Diego Bay Clean-up Sites

Prepared by
Bruce Thompson
Aroon R. Melwani
Jennifer A. Hunt

Aquatic Science Center
for the San Diego Regional Water Quality Control Board

7770 Pardee Lane
Second floor
Oakland, CA 94621
p: 510-746-7334
f: 510-746-7300

This report should be cited as:

Thompson, B., Melwani, A.R., and Hunt, J.A. (2009). Estimated Sediment Contaminant Concentrations Associated with Biological Impacts at San Diego Bay Clean-up Sites, SWRCB Agreement No. 08-194-190, Contribution No. 584, Aquatic Science Center, Oakland, California.

EXECUTIVE SUMMARY

The San Diego Regional Water Quality Control Board (SDRWQCB) recently conducted sediment quality assessments at numerous stations in San Diego Bay using the California Sediment Quality Objectives (SQO) methods. Stations were sampled during four studies conducted between 1998 and 2003. Sediments at many of the stations were considered to be impacted to some degree. However, the SQO methods do not identify the contaminants associated with impacts, or concentration thresholds for impacts.

The objectives of this study were to: 1) identify contaminants that may be adversely impacting biological resources (benthic community, sediment toxicity) in San Diego Bay and, 2) estimate statistical impact limits for sediment contaminants that may be used for future sediment clean-up efforts.

The analyses detailed in this report suggested that biological impacts in San Diego Bay are probably associated with mixtures of sediment contaminants at several spatial scales. Most of the samples contained multiple contaminants that were significantly correlated with the biological indicators. Multivariate analysis showed that most contaminants covaried, and that the covarying sediment mixtures were usually significantly associated with benthic and/or toxic impacts. There was no evidence that any individual contaminant may be responsible for biological impacts.

Impact limits (confidence and prediction limits) for individual contaminants were estimated from pools of samples that were previously assessed as impacted or un-impacted by the SQO assessments. The results presented include sets of numerical impact limits for each contaminant that will provide a range of options for future sediment contamination clean-up efforts. The options include consideration of the spatial scale at which future clean-up levels might be applied; whether the limits are based on the impacted or un-impacted samples; the statistical probability level at which clean-up might be evaluated; and the number of future samples that could be used to assess clean-up progress and success. Other published sediment guidelines and threshold values are included for guidance and context in the selection of appropriate calculated impact limits.

Since sediment contaminant mixtures were always associated with biological impacts, clean-up efforts will also need to consider how to assess remediation of sediment mixtures. Although there are currently no widely used clean-up levels for sediment contamination mixtures, impact limits were calculated for three mixture indicators. These mixture indicator limits could be used, along with individual contaminant limits, to assure that all mixture components are being reduced by the applied clean-up and remediation strategies.

The results presented in this report should not be interpreted as showing cause and effect between contaminants and biological impacts. Only a limited number of contaminants (10) were included in the analyses, and many other contaminants probably exist in San Diego Bay (*e.g.* pyretheroids, PBDEs, etc.), that may also contribute to biological effects in the Bay. The analyses were based on associations within the existing data, and while the results often showed statistical significance, conclusions about specific causes of observed biological impacts should not be made. However, the analytical methods used and the numerical limits presented are well-founded and represent the most rigorous values that can be derived at this time.

Ultimately, there are no correct or 'right' limits to select. The choice will likely be determined during negotiations between regulators and stakeholders, and this report is intended to provide the necessary technical information to facilitate these discussions.

TABLE OF CONTENTS

INTRODUCTION

The State of California has adopted regulatory sediment quality objectives (SQOs) for enclosed bays and estuaries in California. The SQO assessments use multiple lines of evidence based on data for sediment chemistry, sediment toxicity, and benthic community condition (SWRCB 2008). The SQO methods provide an assessment of the magnitude of biological impact related to sediment contamination at a site.

The San Diego Regional Water Quality Control Board (SDRWQCB) recently conducted SQO assessments at numerous sites in San Diego Bay, and the result indicated biological impacts related to sediment contamination in some locations. However, the SQO methods do not identify which contaminants may be associated with impacts or provide concentration thresholds for apparent biological impacts.

The goal of this study was to identify sediment contaminants that were statistically associated with biological impacts in San Diego Bay, and to estimate concentration levels that may provide the SDRWQCB with a basis for future sediment clean-up efforts. The objectives of this study were to:

1. Identify contaminants that may be adversely impacting the benthic community and causing sediment toxicity, and
2. Estimate statistical impact limits for those contaminants.

Strategies to reduce contaminant related impacts to biological resources in San Diego Bay sediments first require the identification of which contaminant(s), at what concentrations, are associated with observed biological impacts. Then, decisions about the appropriate contaminant(s) to focus on, and the appropriate sediment concentration limits to use, can be made. This report provides information and options that may be used to make these decisions.

METHODS

Data

Data used in this report were from previous studies conducted throughout San Diego Bay (n = 161, Figure 1). Eighty-seven of those samples were collected in, or around, four prospective sediment clean-up sites, as reported in previous studies from the NASSCO shipyards, Southwest Marine shipyards, Chollas Creek, and Paleta Creek. Specifically, sediment chemistry, sediment toxicity and benthic macrofaunal data reported in the studies (Bay *et al*. 2000, Exponent 2002, Noblet *et al.* 2003, Ranasinghe *et al*. 2003, Bay *et al*. 2005, SCCWRP and U.S. Navy 2005, Schiff *et al*. 2006, Ranasinghe *et al.* 2007) were used. In addition to the samples collected at the clean-up sites, stations sampled in the Southern California Bight Program (1998 and 2003) were included in analyses to increase samples sizes for identifying contaminant gradients. No new data was collected.

Calculated variables such as control adjusted percent mortality (amphipod toxicity), benthic indices, and sums of organic contaminants were those previously reported. Quality assurance and quality control programs (QA/QC) followed best practices used in Southern California Bight studies and conformed to SQO data quality. It was assumed that the QA/QC for each study was adequate. A check of the comparability of the metals data from each study showed that the lowest detectable concentrations were an order of magnitude or greater than the maximum method detection limits (Table 1). For organic contaminants, this evaluation indicated that maximum detection limits were often higher than the lowest detectable limits. However, this observation was skewed by the summing of many low concentrations, particularly for PCB congeners. These data were considered appropriate and comparable for the correlation and multivariate analyses used in this report.

SQO Assessments. The SQO assessment methods have been described in a series of technical reports and publications available on the websites of the State Water Resources Control Board (SWRCB; www.waterboards.ca.gov) and Southern California Coastal Water Research Project (SCCWRP; www.sccwrp.org). Therefore, SQO methods are only briefly summarized here, as needed for the understanding and interpretation of this report.

The SQO framework uses the sediment quality triad concept of multiple lines of evidence (Long and Chapman 1985, Chapman *et al.* 1997) based on measurements of sediment contamination (chemistry), sediment toxicity tests, and benthic macrofaunal assessments. Summaries for each line of evidence (LOE) are provided below. The three LOEs provide a weight-of-evidence for sediment condition. The LOE category scores are combined into a final SQO site assessment (called the SQO score in this report), and a narrative, in one of five SQO assessment categories (Bay and Weisberg 2007):

SQO Score	Narrative
1	Un-impacted
2	Likely un-impacted
3	Possibly impacted
4	Likely impacted
5	Clearly impacted

A SQO score of 2 is the threshold for the ‘impacted’; categories; scores of 3 or more are considered to indicate increasing levels of impact. The SWRCB Staff Report describes the SQO framework and indicators in further detail (SWRCB 2008, and Appendices A - C).

Sediment Chemistry

Chemicals Evaluated. Sediment contamination was evaluated using 10 commonly collected trace metals and organic compounds (Table 2). These contaminants were selected because they were included in the SQO chemistry line of evidence (LOE) and were measured in the majority of samples included in the studies used for this report. However, some measurements were not included in all of the studies. DDTs and chlordanes were not measured at the NASSCO- or Southwest Marine shipyard areas, and sediment grain-size (percent fines) and total organic carbon were not measured in the majority of samples from the Chollas- or Paleta Creek areas, which limited their inclusion in some analyses.

Sediment Contaminant Mixture Indicators. Numerous publications have shown that mixtures of contaminants can exacerbate biological responses (Swartz *et al.* 1995, Long *et al.* 1998). Therefore, three indices of sediment contamination mixtures were used in this report, in addition to the individual contaminants. Two of the mixture indices were those used in the SQO methods: the California Logistic Regression Model (CA LRM) index indicates the highest probability of sediment toxicity, based on logistic regression models, among the contaminants tested (Bay and Weisberg 2007). The mean Chemistry Score Index (mCSI) is based on benthic responses to mixtures of contaminants (Ritter *et al.* 2008). The most commonly used mixture indicator is the mean Effects Range Median quotient, or mERMq (Long *et al.* 1998). Numerous publications have shown relationships between mERMq and benthic impacts (Hyland *et al.* 2002, Thompson and Lowe 2004) and sediment toxicity (Carr *et al.* 1996, Thompson *et al.* 1999). The mERMq also provides an indicator independent of the SQO methods, and is calculated as follows: the concentration of each individual sediment contaminant is divided by its ERM value (Long *et al*. 1998) to yield an ERM quotient. Then, the quotients for each contaminant are summed, and averaged by the number of contaminants assessed.

Benthic Indicators

The SQO methodology uses four independently derived benthic indices for the assessment of benthic condition. Benthic sampling methods, metrics, and methods for calculation of each index are described in SWRCB (2008) and summarized in Ranasinghe *et al*. (2009). Each of the benthic index scores is assigned to one of four categories of benthic impact, and the median of the four categories is the benthic LOE score (ranging 1 to 4) for a sample. However, the benthic LOE category is not well suited for the determination of statistical relationships between benthic condition and contamination, because the four category values create a limited range of variation for comparison to sediment contaminant concentrations.

Mean Benthic Index. A mean benthic index was calculated based on the scores of the four benthic indices used in the SQO assessments. The four benthic index scores were scaled to approximately 0-100 (Benthic Response Index, BRI, was already scaled), and the mean of the four scores yielded the mean benthic index. Comparing the mean benthic index to the benthic LOE score showed that mean benthic index values above 55 were always associated with benthic impacts (score = 3 or 4). Therefore, a mean benthic index value of 55 appears represents a benthic impact limit for San Diego Bay.

Sediment Toxicity Indicators

Sediment toxicity was measured using the 96-h *Eohaustorius estuarius* (amphipod) test (ASTM 1992). It is a static renewal test, using field collected sediments in the laboratory. Control-adjusted percent survival was the endpoint used in the analyses, and is the same metric used to determine the SQO toxicity LOE in the studies included herein.

Statistical Analysis

Spatial Scales. Conceptually, the most accurate associations between biological impacts and sediment contamination will be derived from specific locations exhibiting a clear sediment contamination gradient. When samples from large areas are pooled, several gradients of contaminants from several sources may be combined and any biological response to a specific contaminant may be confounded. Individual contaminant gradients are optimal for determining impact thresholds, but these rarely exist in urban estuaries. Identifying locations with gradients of sediment contamination that are statistically associated with biological impacts may present the best opportunity to determine impact limits (described below) in San Diego Bay.

Analyses were conducted at three spatial scales: 1) all San Diego Bay stations (n = 161) combined, 2) four prospective sediment clean-up sites combined (Southwest Marine shipyards, NASSCO shipyards, Chollas Creek, and Paleta Creek; n = 87), and 3) samples from each the four clean-up sites (site-specific). Samples included at each spatial scale are shown in Figure 1. Appendix I also summarize mean and standard deviations of the biological indicators at the two larger spatial scales.

Correlation Analysis. Correlations (Spearman's rank) between biological indicators (mean benthic index, control adjusted amphipod percent survival) and sediment contamination were used to evaluate basic associations at the three spatial scales. Pearson's correlations were evaluated using both raw and transformed variables (log10 or arc-sin) prior to inclusion in multivariate analysis. Transformations were necessary to create as much linearity as possible in the variables, for use in subsequent multivariate analyses that are based on linear models.

Multivariate Analysis. Principal components analysis (PCA) was used to evaluate the general sediment contamination patterns in San Diego Bay, and to identify co-occurring suites of contaminants that were statistically related to biological impacts. These analyses were conducted at three spatial scales, in two steps:

1. PCA was conducted using only contaminants that were significantly correlated with the biological indicator (mean benthic index or amphipod percent survival), selecting the raw or transformed variables with the highest significant Pearson coefficients. Factors with eigenvalues greater than 1 were retained, and where more than one factor was retained, varimax rotation was applied. The resulting PCA factor(s) represent independent sets of co-varying contaminants that were used as independent variables in the multiple regression analysis. The PCA factors represent the maximum separation of individual contaminants possible using co-occurrence data.

2. Multiple regression analysis was used to determine which PCA factors (independent variables) were statistically associated with each biological indicator (dependent variable). Where more than one PCA factor was included, the stepwise variable selection method was applied to determine the independent variables that provided the best 'model' for each biological indicator. Although sediment type (grain-size and TOC) also influence biological responses, these variables were not included in the analyses because they were not sampled at all sites, and because the primary goal of the multivariate analysis was to evaluate contaminant influence on impacts in order to identify contaminants on which to focus remediation. Partial regression coefficients were used to evaluate the relative contributions of the independent variables.

Impact Limits. Statistical limits were calculated separately for stations categorized in the SQO assessment as impacted (SQO scores >= 3) and un-impacted (SQO scores <=2). These limits were determined at the three spatial scales, similar to those used in correlation and multivariate analyses.

Confidence limits were calculated to show the expected range of the mean concentration based on the current data, for a chosen level of statistical probability (*e.g.,* 95%). Prediction limits are similar to confidence limits, except that they indicate the expected range of the mean in any future surveys of a chosen sample size (n). Confidence limits and prediction limits were calculated for levels of 95%, 90%, and 80% statistical probability. As examples of how the prediction limits could be used, they were calculated based on future sample sizes of n = 3 and n = 5. For this study, one-tailed upper statistical limits of un-impacted samples, and one-tailed lower statistical limits of impacted samples were calculated. Both confidence-, and prediction limits were calculated as options for consideration in determining clean-up levels at several scales. Thus, in this report, the term 'impact limit' is used to refer to these values that are empirically derived statistical limits for apparent biological impacts as defined by the SQO assessments.

An estimated benthic impact limit was determined for several individual contaminants. It was determined graphically from plots the mean benthic index and individual contaminant concentrations, and fitting a smoothed curve to the data (*e.g.,* Figure 2; SASGraph interpolation method: sm60). The estimated limit was identified as the point on the x-axis that corresponded to where the curve passed above a mean benthic index value of 55 (the index value above which all samples were 'impacted').

Interpretation. The determination of which contaminants may be causing biological impacts can only be accomplished through controlled experimentation, and no such experiments have been conducted in San Diego Bay. Currently, the only way to identify possible causative contaminants is through numerical analysis, such as those conducted in this report. However, there are many limitations to this type of analysis. This study used existing data with a limited number of contaminants, and correlation methods to associate those data with co-occurrences of toxicity and benthic impacts. While the results provide estimates of probability within the data, one must also consider the following in interpretation: 1.) Many other contaminants and abiotic factors that were not included in the analyses can influence benthic communities and toxicity test results. 2.) The toxicology of mixtures of contaminants is poorly understood. 3.) Correlation analyses do not demonstrate causality, only statistical association. Although not perfect, it is important to conduct these analyses to begin to understand relationships between contaminants and impacts.

RESULTS AND DISCUSSION

Sediment contamination occurred throughout San Diego Bay as mixtures of all 10 contaminants evaluated. When all San Diego Bay sites were pooled, principal components analysis (PCA) identified two independent factors (Table 3). Five trace organic contaminants (PCBs, DDTs, HPAHs, LPAHs, chlordanes) and cadmium composed factor 1, and four trace metals (copper, mercury, lead, and zinc) composed factor 2. When the pooled sediment clean-up area samples were analyzed, a similar pattern was observed, but lead and copper were also included on factor 1. The similar results at two different spatial scales demonstrate that mixtures of contaminants predominate in sediments of the Bay; no single contaminant formed a PCA factor. The sources and physical processes that produced the two independent patterns were not further evaluated.

Relationships between Sediment Chemistry, Benthos and Toxicity

Most of the 10 sediment contaminants evaluated were significantly correlated with percent amphipod survival, the SQO toxicity LOE category, the SQO benthic LOE category, mean benthic index, and the SQO score, when all San Diego Bay samples were pooled (Table 4). The only non-significant correlations were with the benthic LOE for mercury and TOC. The mean benthic index usually had the highest correlations with the sediment variables, but the SQO score was most highly correlated with two of the mixtures indicators, mERMq and CA LRM. Zinc was consistently one of the individual contaminants most highly correlated with all biological indicators. Sediment type (percent fines and TOC) was significantly correlated with all biological and SQO indicators except for the benthic LOE.

Using the pooled samples from the sediment clean-up areas, chlordanes and DDTs had the highest correlations with all biological and SQO indicators. However, these contaminants were not measured at the two shipyard areas (NASSCO and Southwest Marine). Copper, mercury, percent fines, and TOC were often not significantly correlated with the toxicity or benthic indicators. Relationships for mean benthic index and toxicity are illustrated for each contaminant in Figures 3, 5, 7, 9, 11, 13, 15, 17, 19, and 21).

The three mixture indicators were highly inter-correlated in San Diego Bay (Spearman's r > 0.81, $p < 0.0001$, $n = 161$). Correlations between the three indicators of sediment mixtures and the biological and SQO indicators were usually among the highest correlations when all Bay samples were pooled. However, correlations between chlordanes and DDTs and the biological and SQO indicators were highest in the clean-up areas. The correlation analysis results provide context for the following analyses, and should not be interpreted alone related to biological impacts. As shown above, the sediments in the Bay were composed of mixtures of contaminants, and the relationships of these mixtures with biological impacts are more properly evaluated using multivariate analyses.

The PCA analysis identified sub-sets (factors) of co-varying sediment contaminants (mixtures). These factors represent the maximum degree of separation of individual contaminants possible in correlation analysis, and appropriately account for co-variation among the individual contaminants in sediment mixtures. Multiple regression analysis showed the statistical relationships between these factors (mixtures) and the mean benthic index and percent amphipod survival. When all San Diego Bay samples were pooled, two PCA factors, composed of all 10 contaminants, contributed to a significant multiple regression model that accounted for approximately 44% of the variation in the mean benthic index, and approximately 26% of the variation in amphipod survival (Table 5). The R^2 value was moderate because there are numerous other chemical and physical factors that can affect the benthos, including other unmeasured sediment contaminants, which were not included in the regression model. Additionally, the R^2 value may reflect confounding of any site-specific benthic responses by pooling many stations with differing contamination gradients and responses. A similar analysis with amphipod toxicity did not generate a significant regression model, and showed a weak association between sediment contamination and toxicity in San Diego Bay sediments.

Analyses using only the pooled clean-up area stations produced similar results to those described above, but only single PCA factors were identified (Table 5). Mercury was not significantly correlated with the mean benthic index, and mercury, copper and PCBs were not significantly correlated with toxicity in the clean-up areas, and thus were not included in the analysis. The single PCA factor shown at the clean-up area scale indicates that sediments in these areas are well mixed, with no obvious pattern in space or time. The single PCA factor created significant regression models that accounted for approximately 41% of the variation in mean benthic index, and 33% of the variation in amphipod toxicity.

Analyses at each of the four sediment clean-up sites identified single PCA factors composed of between five and nine individual contaminants, except for toxicity at Chollas Creek where two factors were identified (Table 5). At the Southwest shipyard, all five metals and three organic contaminants contributed to a significant regression model that accounted for nearly 85% of the variability. However, only HPAHs was significantly correlated with toxicity at the Southwest shipyard, and multivariate analysis was not conducted. At the NASSCO shipyard, five contaminants formed a single factor that was significantly related to benthic impacts, but none of the contaminants were significantly correlated to amphipod toxicity. At Chollas Creek, five contaminants contributed to one (benthos), or two (toxicity) factors that were significantly related to impacts. At Paleta Creek, nine contaminants formed a single factor significantly related to benthos, and seven contaminants formed a single factor significantly related to toxicity. However, less than 20% of the variation was accounted for in the toxicity regression model.

Slightly different sets of contaminants were associated with impacts at the individual sediment clean-up sites, probably reflecting differences in sources of the contaminants (shipyards, stormwater inputs at the creeks, etc). However, the results always associated sediment mixtures, and not individual contaminants. All of the individual contaminants evaluated contributed to the mixtures, and were significantly associated with impacts at one or another of the clean-up sites (Table 6). All contaminants were included in the mixtures at the San Diego Bay scale, and all except mercury were included at the pooled clean-up area scale. PCBs, copper, mercury, and lead were not associated with impacts at Chollas Creek, and copper, mercury, and zinc were not associated with impacts at NASSCO.

Impact Limits for Individual Contaminants

Statistical confidence limits and prediction limits (called impact limits in this report) were calculated from the pool of stations characterized as impacted (SQO score > = 3) or un-impacted (SQO score <= 2) by the SQO assessments previously conducted. Upper limits of the pooled un-impacted stations identify the highest concentration below which samples are representative of the population of un-impacted stations at three levels of statistical probability (Tables 8 - 17). Conversely, lower limits were derived from the impacted stations, and identify the lowest concentration above which samples are representative of the population of impacted stations at the three levels of probability. The limits provide options for contamination clean-up levels based on limits for un-impacted samples, or limits for impacted samples. Generally, contaminants with smaller sample sizes (*e.g.,* DDTs and chlordanes), or large differences in the mean of un-impacted samples compared to impacted samples (*e.g.,* HPAHs) had wider limits relative to the mean (Table 7). The selection of appropriate impact limits for future clean-up efforts will require consideration of a number of factors:

A. Scale of Application. A major factor in determining the appropriate impact limits is the scale at which future clean-up levels will be applied. Impact limits were calculated at three spatial scales: 1) all San Diego Bay, 2) the pooled sediment clean-up areas (NASSCO, Southwest Marine, Chollas Creek, and Paleta Creek), and 3) pools of samples selected from the clean-up sites (site-specific) that commonly identified specific contaminants as components of mixtures that were specifically associated with impacts, as shown on Table 6. For example, limits for copper were calculated using samples from Southwest and Chollas Creek because copper was among the contaminants that were components of the PCA factor that was significantly associated with benthic impacts at those areas. But cadmium was a mixture component at all four clean-up sites so no site specific limits were calculated; instead, limits using pooled samples from all four clean-up sites were calculated. Mean contaminant values at each spatial scale are shown on Table 7, and are useful when evaluating the impact limits.

If impact limits are to be applied uniformly throughout San Diego Bay, then limits derived from all San Diego Bay stations may be considered. If limits are to be used at all of the sediment clean-up areas, then the values derived from the pool of these samples may be considered. The impact limits for individual contaminants from either the pooled clean-up areas, or the selected clean-up sites (site-specific), as indicated on Table 6, should provide the most accurate and rigorous impact limit estimates. However, sample size affects the utility of these values, as discussed below.

B. Type of Limits. Statistical confidence limits and probability limits are presented on Tables 8 – 17. Confidence limits are generally used to determine whether mean concentrations based on current data will be within a statistical range. Confidence limits at a chosen probability level (*e.g.,* 95%) are commonly selected for statistical interpretations, but may not be appropriate or attainable when selecting realistic clean-up levels. Differing from confidence limits, prediction limits are commonly used in water quality monitoring to determine whether mean concentrations in future sampling efforts will be within a chosen limit. The upper prediction limits estimated from un-impacted samples could be employed in such a way. The alternative scenario, using the lower prediction limit of impacted samples may also be useful. The choice of whether to use limits based on impacted (lower limits) or un-impacted (upper limits) samples will influence how attainment of future clean-up levels will be identified, thus should be considered carefully. In general, the lower confidence limits calculation resulted in higher concentrations than the corresponding prediction limit. Conversely, upper confidence limits were generally lower in concentration than the corresponding prediction limit. This is probably due to differences in sample size (n) effects on the variation. In many instances, the lower limits of both confidence and prediction limits corresponded to higher concentrations than the upper limits. This was due to the upper and lower limits being calculated using distinct components of the dataset, each with their own associated variability estimates. Therefore, either the upper- or lower limit could be chosen for clean-up rather than using both, since concentrations that fall between limits would be confusing to classify.

An additional consideration for the use of prediction limits is that the calculation includes a term for future sample size. Therefore, prediction limits can be calculated for any sample-size that may be used to assess clean-up efforts and recovery. However, the number of future samples should be selected *a priori* to calculate clean-up levels. Prediction limits based on future sample sizes of n = 3 and n = 5 were calculated in this study to demonstrate the influence of sample size choices on concentration limits (Tables 8 – 17). Generally, the higher sample size resulted in lower limits because the calculations are based on standard errors which are sample size dependent.

C. Probability Levels. Prediction limits have been presented graphically to show the influence of statistical probability (95%, 90%, and 80%) choices on calculated values (Figures 4, 6, 8, 10, 12, 14, 16, 18, 20, and 22). Generally, the lower limits at 95% probability levels lower concentrations (more conservative) than those calculated for 90% or 80%. Conversely, the upper prediction limits resulted in higher concentrations at 95%. Overall, this pattern resulted in larger ranges between the lower and upper limits at the 95% probability level, compared to the limits at 90% and 80% probability. This would indicate that the decision of whether to apply either the lower or upper limits will have less of an influence on results at levels of 80% and 90%, but a significant influence at the 95% level.

Assuming similar variability in chemical concentrations in future samples, the cumulative distribution functions (CDFs; even numbered Figures 4-22) should also be considered when determining appropriate statistical probability levels. For example, selecting the 95% upper prediction limit for HPAHs would result in approximately 60% of samples being representative of un-impacted conditions (Figure 19). If lower probability levels are acceptable, then values associated with 80% statistical probability may be chosen. Selection of the 80% upper prediction limit for HPAHs would results in approximately 35% of future samples being un-impacted. However, as discussed above, these impact limits are only presented as examples. Ultimately, impact limits should be derived based on expected future sample designs in order to provide meaningful values for interpretation.

D. Comparison to Other Effects Thresholds. The efficacy of the impact limits presented above may be evaluated by comparison to other existing biological effects guidelines and limits to provide some context for the chosen levels (Table 18). Effects Range Low (ERL) and Effects Range Median (ERM) values have been widely used to evaluate sediment contamination levels in both scientific and policy settings (Hunt *et al*. 2001, Long *et al.* 2006). In general, impact limits corresponded to concentrations between the ERL and ERM. For cadmium, lead, DDTs, LPAHs, and HPAHs, impact limits were always below the ERL. Upper impact limits exceeded the ERM in a few cases, notably PCBs, copper, mercury, and chlordanes.

LC50 values from sediment toxicity studies provide other sets of limits, and past studies in San Diego Bay have derived lowest observable effects levels and prediction limits as potential future clean-up levels (*e.g.,* Exponent *et al*. 2002). As evident from Table 18, these effects levels varied greatly relative to the impact limits, reflecting the differences in methods for calculating relevant effects levels. Estimated benthic impact limits based on the data used in this study are also included for comparison.

Impact Limits for Sediment Contamination Mixtures

All of the analyses presented in this report showed that mixtures of contaminants were significantly associated with biological impacts at all scales evaluated. While it may be necessary to apply some, or all of the individual impact limits presented above for clean-up, it is recommended that limits for sediment contamination mixtures also be used. Although the mechanisms of biological impacts from multiple contaminants are poorly understood, and biological thresholds for mixtures have generally not been determined, using statistical impact limits for mixtures, along with individual contaminant limits, would provide consistency with the results of the analyses presented that show the pervasiveness of mixtures.

Impact limits were calculated for three sediment contamination mixtures indicators. All three of these indicators were highly inter-correlated in the region, generally reflecting similar mixture gradients. For mERMq, the upper 95% confidence limit for un-impacted samples in San Diego Bay combined was 0.191 (Table 19). Prediction limits provided higher limits. For context, benthic impact ranges for mERMq published elsewhere were 0.147 in San Francisco

Estuary (Thompson *et al*. 2004), and 0.036 on the Atlantic and Gulf coasts (Hyland *et al.* 2002). In this study, San Diego Bay samples with mERMq below 0.054 were never impacted, and samples above 1.362 were always impacted. Samples between these limits showed increasing incidence of impacts with increasing mERMq values. It should be noted that mERMq values were calculated using more than the 10 contaminants included in this study, as recommended in the SQO methods, thus, the mERMq results reflect a broader mixture of contaminants, possibly providing increased protection for biological resources.

For mCSI, the upper 95% prediction limits (n=5) for un-impacted samples in the Bay was 2.387 (Table 20). This value is very close to the mCSI "moderate exposure" (category 3) lower limit of 2.33, used in the SQO assessments conducted previously. Many of the impact limits calculated for mCSI were near the SQO value, demonstrating good consistency with the SQO methods. Samples with mCSI values below 1.14 were never impacted, and samples with values above 3.14 were always impacted.

For the CA LRM, the upper 95% prediction limits (n=5) for un-impacted samples in San Diego Bay was 0.621 (Table 21). This value is well above the LRM "moderate exposure" (category 3) lower limit of 0.490, as are all the calculated impact limits. For additional context, CA LRM values below 0.389 were never impacted, and values above 0.875 were always impacted.

CONCLUSIONS

San Diego Bay sediment samples contained multiple contaminants that were significantly associated with biological impacts at all spatial scales evaluated. Benthic indicators generally had stronger associations with sediment contamination mixtures than sediment toxicity. Most of the individual contaminants were significantly correlated with the biological indicators, but there were some exceptions at the clean-up area and site-specific scales. Specifically, only HPAHs were correlated with amphipod survival at Southwest Marine shipyards, and no contaminants were correlated with amphipod survival at NASSCO shipyards. Despite significant correlations between individual contaminants and biological impacts, no individual contaminant could be separated from the mixtures in the multivariate analyses conducted. However, the components of the mixtures associated with impacts varied slightly at each of the four clean-up sites, and some contaminants may be ruled-out based on a lack of significant correlations with biological impacts at specific locations.

Mainly because of the limited number of contaminants included in this report, and the correlational nature of the analyses presented, conclusions regarding the specific causes of observed biological impacts cannot be made. Many other contaminants probably exist in San Diego Bay, but were not measured in the studies included in this report. Additionally, numerous pesticides (*e.g.,* pyretheroids), brominated trace organics (PBDEs), and other kinds of contaminants may also contribute to biological impacts in the Bay.

The impact limits presented in this study provide a range of options for consideration in choosing sediment contaminant concentration limits for clean-up and restoration. These options include consideration of whether the limits are based on the impacted or un-impacted samples, the spatial scale at which future clean-up levels might be applied, the statistical probability level at which future success might be evaluated; and the number of future samples that could be used to assess clean-up success. Impact limits for the pooled clean-up areas, or selected clean-up sites should provide the most rigorous and accurate values. Impact limits are not true biological thresholds, but statistical values calculated from existing data; no cause and effect should be concluded. At this time these values are the best estimates of reasonable clean-up levels. Future clean-up efforts should also address mixtures of sediment contaminants. Impact limits for three sediment contaminant mixture indices were also included in this report.

Ultimately, there are no correct or 'right' limits to select. The choice will likely be determined during negotiations between regulators and stakeholders, and this report is intended to provide the necessary technical information to facilitate these discussions.

ACKNOWLEDGMENTS

The Southern California Coastal Water Research Project (SCCWRP) provided the SQO calculation and benthic indicator results used in the report. Discussions with Steve Bay at SCCWRP helped clarify and focus the results presented in this report.

LITERATURE CITED

Anderson, B. A., S. Lowe, J. W. Hunt, B. M. Phillips, J. P. Voorhees, S. L. Clark, and R. S. Tjeerdema. 2008. Relative sensitivities of toxicity test protocols with the amphipods *Eohaustorius estuarius* and *Ampelisca abdita*. Ecotoxicology and Environmental Safety 69:24-31.

ASTM. 1992. Designation E 1367: standard guide for conducting 10-day static sediment toxicity tests with marine and estuarine amphipods. American Society for Testing and Materials, Philadelphia, PA.

Bay, S. M., D. Lapota, J. Anderson, J. Armstrong, T. Mikel, A. W. Jirik, and S. Asato. 2000. Southern California Bight 1998 Regional Monitoring Program: IV. Sediment toxicity. Technical Report # 339, Southern California Coastal Water Research Project, Westminster, CA.

Bay, S. M., T. Mikel, K. Schiff, S. Mathison, B. Hester, D. Young, and D. Greenstein. 2005. Southern California Bight 2003 Regional Monitoring Program: I. Sediment Toxicity. Technical Report # 451, Southern California Coastal Water Research Project, Westminster, CA.

Bay, S. M., and S. B. Weisberg. 2007. A framework for interpreting sediment quality triad data SCCWRP Draft Technical Report, Costa Mesa, CA. 23 pp.

Carr, R. S., E. R. Long, H. L. Windom, D. C. Chapman, G. Thursby, G. M. Sloane, and D. A. Wolfe. 1996. Sediment quality assessment studies of Tampa Bay, Florida. Environmental Toxicology and Chemistry 15:1218-1231.

Chapman, P. M., B. Anderson, S. Carr, V. Engle, R. Green, J. Hameedi, M. Harmon, P. Haverland, J. Hyland, C. Ingersoll, E. Long, J. Rodgers, M. Salazar, P. K. Sibley, P. J. Smith, R. C. Swartz, B. Thompson, and H. Windom. 1997. General guidelines for using the sediment quality triad. Marine Pollution Bulletin 34:368-372.

Exponent. 2002. Phase 1 bioaccumulation data, ecological receptor species, and receptor parameters for the NASSCO and Southwest Marine detailed sediment investigation. Exponent, Bellevue, WA.

Hunt, J. W., B. S. Anderson, B. M. Phillips, J. Newman, R. S. Tjeerdema, R. Fairey, H. M. Puckett, M. Stephenson, R. W. Smith, C. J. Wilson, and K. Taberski. 2001. Evaluation and use of sediment quality toxicity reference sites for statistical comparisons in regional assessments. Environmental Toxicology and Chemistry 20:1266-1275.

Hyland, J. L., W. L. Balthis, V. D. Engle, E. R. Long, J. F. Paul, J. K. Summers, and R. F. Van Dolah. 2002. Incidence of stress in benthic communities along the US Atlantic and Gulf of Mexico coasts within different ranges of sediment contamination from chemical mixtures. Environmental Monitoring and Assessment 81:149-161.

Long, E. R., and P. M. Chapman. 1985. A sediment quality triad: Measures of sediment contamination, toxicity, and infaunal community composition in Puget Sound. Marine Pollution Bulletin 16:405-415.

Long, E. R., L. J. Field, and D. D. MacDonald. 1998. Predicting toxicity in marine sediments with numerical sediment quality guidelines. Environmental Toxicology and Chemistry 17:714-727.

Long, E. R., C. G. Ingersoll, and D. D. MacDonald. 2006. Calculation and uses of mean sediment quality guideline quotients: A critical review. Environmental Science and Technology 40:1726-1736.

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.

Noblet, J. A., E. Y. Zeng, R. Baird, R. W. Gossett, R. J. Ozretich, and C. R. Phillips. 2003. Southern California Bight 1998 Regional Monitoring Program: VI. Sediment chemistry. Technical Report # 381, Southern California Coastal Water Research Project, Westminster, CA.

Ranasinghe, A., S. B. Weisberg, R. W. Smith, D. E. Montagne, B. Thompson, J. Oakden, D. D. Huff, D. B. Cadien, R. G. Velarde, and K. Ritter. 2009. Calibration and evaluation of five indicators of benthic community condition in two California bay and estuary habitats. Marine Pollution Bulletin 59:5-13.

Ranasinghe, J. A., A. M. Barnett, K. Schiff, D. E. Montagne, C. Brantley, C. Beegan, D. B. Cadien, C. Cash, G. B. Deets, D. R. Diener, T. K. Mikel, R. W. Smith, R. G. Velarde, S. D. Watts, and S. B. Weisberg. 2007. Southern California Bight 2003 Regional Monitoring Program: III. Benthic Macrofauna. Technical Report # 529, Southern California Coastal Water Research Project, Costa Mesa, CA.

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. Technical Report # 382, Southern California Coastal Water Research Project, Westminster, CA.

Ritter, K., S. M. Bay, R. W. Smith, D. E. Vidal-Dorsch, and L. J. Field. 2008. Development and evaluation of sediment quality guidelines based on benthic macrofauna responses. 2008 Annual Report, Southern California Coastal Water Research Project, Costa Mesa, CA.

SCCWRP, and U.S. Navy. 2005. Sediment assessment study for the mouths of Chollas and Paleta Creek, San Diego. Prepared by the Southern California Coastal Water Research Project and the Space and Naval Warfare Systems Center, U.S. Navy.

Schiff, K., K. Maruya, and K. Christenson. 2006. Southern California Bight 2003 Regional Monitoring Program: II. Sediment Chemistry. Technical Report # 492, Southern California Coastal Water Research Project, Westminster, CA.

Swartz, R. C., F. A. Cole, J. O. Lamberson, F. S.P., D. W. Schultz, W. A. DeBen, H. Lee II, and R. J. Ozretich. 1994. Sediment toxicity, contamination and amphipod abundance at a DDT- and Dieldrin-contaminated site in San Francisco Bay. Environ Toxicol Chem 13:949-962.

Swartz, R. C., Schults, D.W., Ozretich, R.J., Lamberson, J.D., Cole, F.A., DeWitt, T.H., Redmond, M.S., Ferraro, S.P. 1995. Sum PAH: a model to predict the toxicity of polynuclear aromatic hydrocarbon mixtures in field-collected sediments. Environmental Toxicology and Chemistry 14:1977-1987.

SWRCB. 2008. Draft Water Quality Control Plan for Enclosed Bays and Estuaries - Part 1 Sediment Quality California State Water Resources Control Board, Sacramento, CA. 35 pp.

Thompson, B., B. Anderson, J. Hunt, K. Taberski, and B. Phillips. 1999. Relationships between sediment contamination and toxicity in San Francisco Bay. Marine Environmental Research 48:285-309.

Thompson, B., and S. Lowe. 2004. Assessment of macrobenthos response to sediment contamination in the San Francisco Estuary, California, USA. Environ Toxicol Chem 23:2178-2187.

Weston, D. P. 1996. Further development of a chronic *Ampelisca abdita* bioassay as an indicator of sediment toxicity: summary and conclusions. Regional Monitoring Program for Trace Substances Annual Report, San Francisco Estuary Institute, Oakland, CA. 108–115 pp.

TABLES

Table 1. Evaluation of comparability among data included in this study. Trace metal concentrations are μ/g; trace contaminant concentrations are ng/g.

	Southern California Bight Program 1998 Bay *et al.* (2000); Noblet *et al.* (2003); Ranasinghe *et al.* (2003)		**Bight 2003** Bay *et al.* 2005; Schiff *et al.* (2006); Ranasinghe *et al.* (2007)		**NASSCO and Southwest Marine shipyards** Exponent (2002)		**Chollas and Paleta Creeks** SCCWRP and U.S. Navy (2005)	
Contaminant	**Minimum Detectable Conc.**	**Maximum MDL**	**Minimum Detectable Conc.**	**Maximum MDL**	**Minimum Detectable Conc.**	**Maximum MDL**	**Minimum Detectable Conc.**	**Maximum MDL**
Cadmium	0.02	0.5	0.01	0.025	0.1	na	0.02	0.4
Copper	18	2	22	0.03	47	na	22	0.24
Lead	12	5	6.7	0.142	21	na	11	1
Mercury	0.06	0.03	0.09	0.005	0.24	na	0.07	0.01
Zinc	43	4	33	0.05	93	na	89	1
Chlordanes	0.64	1	0.2	5.7	nm	na	0.2	12
DDTs	0.9	1	0.2	11	nm	na	1.4	55
LPAHs	46	47	9.9	39	7.8	na	11	22
HPAHs	39	41	10	57	103	na	92	24
PCBs	9.6	9.6	0.6	3.1	24	na	5.3	33

MDL – Method Detection Limit; nm = not measured; na – not available

Table 2. Sediment contaminants evaluated in this study

Metals	Organic Compounds
Cadmium (μ/g)	Chlordanes, total (ng/g)
Copper (μ/g)	DDTs, total (ng/g)
Lead (μ/g)	PCBs, total (ng/g)
Mercury (μ/g)	HPAHs, total (ng/g)
Zinc (μ/g)	LPAHs, total (ng/g)

Table 3. Results of PCA of sediment contaminants in San Diego Bay and the pooled clean-up areas. Concentrations are listed in order of rotated factor loading scores.

All San Diego Bay (n=128)		Clean-up Areas Only (n=57)	
Factor 1	*Factor 2*	*Factor 1*	*Factor 2*
Chlordane	Copper	Chlordanes	Mercury
DDTs	Mercury	DDTs	Copper
HPAHs	Zinc	HPAHs	
PCBs	Lead	LPAHs	
LPAHs		Cadmium	
Cadmium		PCBs	
		Lead	
		Zinc	

Table 4. Spearman's rank correlation coefficients between sediment variables and biological and SQO indicators. Shading = significant coefficient (p<0.05).

Variable	n	Amphipod % Survival	Toxicity LOE	Benthic LOE	Mean Benthic Index	SQO Score
All San Diego Bay						
Cadmium	161	-0.276	0.219	0.324	0.444	0.398
Chlordane	128	-0.298	0.300	0.493	0.517	0.552
Copper	161	-0.375	0.301	0.314	0.547	0.542
DDTs	129	-0.283	0.258	0.535	0.577	0.547
HPAH	161	-0.222	0.172	0.346	0.603	0.466
Lead	161	-0.396	0.317	0.408	0.681	0.607
LPAH	161	-0.251	0.224	0.295	0.477	0.416
Mercury	161	-0.268	0.194	0.084	0.355	0.313
PCBs	161	-0.158	0.061	0.269	0.579	0.393
Zinc	161	-0.463	0.383	0.503	0.706	0.680
Fines	123	-0.469	0.391	0.292	0.559	0.530
TOC	107	-0.310	0.225	0.185	0.624	0.469
mERMq	161	-0.451	0.393	0.413	0.598	0.640
mCSI	161	-0.425	0.343	0.460	0.722	0.655
CA LRM	161	-0.474	0.402	0.508	0.708	0.714
Sediment Clean-up Areas						
Cadmium	87	-0.373	0.334	0.493	0.458	0.516
Chlordane	57	-0.583	0.605	0.575	0.574	0.776
Copper	87	-0.103	0.032	-0.048	0.290	0.229
DDTs	58	-0.491	0.502	0.559	0.582	0.751
HPAH	87	-0.335	0.274	0.241	0.473	0.469
Lead	87	-0.336	0.269	0.171	0.479	0.451
LPAH	87	-0.372	0.310	0.250	0.451	0.486
Mercury	87	-0.042	-0.015	-0.237	0.151	0.045
PCBs	87	-0.252	0.175	-0.014	0.372	0.317
Zinc	87	-0.328	0.270	0.279	0.478	0.473
Fines	52	-0.252	0.184	0.066	0.462	0.352
TOC	39	-0.156	-0.002	0.081	0.503	0.309
mERMq	86	-0.475	0.445	0.360	0.472	0.628
mCSi	86	-0.412	0.368	0.165	0.462	0.496
CA LRM	86	-0.432	0.373	0.340	0.498	0.593

Table 5. Results of PCA and multiple regression analysis showing sediment components that were most highly associated with mean benthic index and amphipod survival, at several spatial scales.

Area	n	Factors and Components	Partial R^2	Model R^2
Mean Benthic Index				
All San Diego Bay	128	F1= cadmium chlordanes DDTs LPAHs HPAHs PCBs F2= copper lead mercury zinc	0.315* 0.124*	0.439 *
Clean-up Areas	56	F1= copper lead HPAHs LPAHs PCBs zinc cadmium chlordanes DDT		0.407*
Southwest	17	F1= lead cadmium copper PCBs LPAHs HPAH mercury zinc		0.849*
NASSCO	14	F1= cadmium lead HPAHs LPAHs PCB		0.677*
Chollas Creek	23	F1= cadmium chlordanes DDTs HPAHs LPAHs		0.500*
Paleta Creek	31	F1= zinc HPAHs PCBs lead LPAHs DDTs copper mercury chlordanes		0.335*
Percent Amphipod Survival				
All San Diego Bay	128	F1= cadmium DDTs chlordanes LPAHs HPAHs PCBs F2= copper lead mercury zinc	0.138 0.126	0.264
Clean-up Areas	56	F1= HPAH LPAH lead zinc DDTs chlordane cadmium		0.325*
Southwest	17	Only HPAHs was significantly correlated with % survival		
NASSCO	14	No contaminants significantly correlated with % survival		
Chollas Creek	23	F1= chlordanes DDTs LPAHs F2= cadmium zinc	0.201* 0.190*	0.392*
Paleta Creek	31	F1= zinc PCBs lead LPAHs cadmium mercury HPAHs		0.192*

Table 6. Sediment contaminants that were significantly correlated with mean benthic index (B) or amphipod toxicity (T), and components of significant multiple regression factors. All contaminants were included when all San Diego Bay samples were pooled. No contaminants were included at Southwest or NASSCO for toxicity; nm = not measured.

Location	n	Cadmium	Copper	Mercury	Lead	Zinc	Chlordanes	DDTs	HPAHs	LPAHs	PCBs
Clean-up Areas	56	B, T	B		B, T	B, T	B, T	B, T	B, T	B, T	B
Southwest	17	B	B	B	B	B	nm	nm	B	B	B
NASSCO	14	B			B		nm	nm	B	B	B
Paleta Ck.	31	T	B	B, T	B, T	B, T	B	B	B, T	B, T	B, T
Chollas Ck.	23	B, T				T	B, T	B, T	B	B, T	

Table 7. Sample size and mean concentrations for samples characterized as either un-impacted (SQO score<3) or impacted (SQO score > 2).

		SQO score < 3		SQO score > 2	
Spatial Scale	**Variable**	**N**	**Mean**	**N**	**Mean**
All San Diego Bay	Cadmium	72	0.209	89	0.342
	Chlordane	57	1.393	71	4.995
	Copper	72	77.302	89	153.159
	DDTs	57	2.362	72	7.302
	HPAH	72	407.126	89	1234.824
	Lead	72	30.528	89	62.710
	LPAH	72	99.0169	89	202.978
	Mercury	72	0.314	89	0.489
	mCSI	72	1.848	89	2.608
	mERMq	72	0.166	89	0.332
	PCBs	72	19.720	89	58.680
	CA LRM	72	0.494	89	0.681
	Zinc	72	136.380	89	256.448
Sediment Clean-up Areas	Cadmium	30	0.210	57	0.494
	Chlordane	18	1.181	39	12.424
	Copper	30	109.484	57	156.399
	DDTs	18	2.564	40	16.408
	HPAH	30	729.323	57	2996.320
	Lead	30	45.020	57	76.907
	LPAH	30	120.897	57	367.160
	Mercury	30	0.417	57	0.467
	mCSI	30	2.124	57	2.858
	mERMq	30	0.211	57	0.443
	PCBs	30	62.443	57	168.519
	CA LRM	30	0.549	57	0.717
	Zinc	30	179.613	57	287.512
Site-specific*	Copper	23	110.862	26	171.283
	Lead	26	45.352	37	76.522
	Mercury	23	0.444	26	0.542
.	PCBs	26	74.993	37	191.988
	Zinc	27	170.063	46	292.429

* Figure 1 and Table 6 identify the stations included at the site-specific scale

Table 8. Impact limits for cadmium. Site-specific limits were not calculated because cadmium was indicated at all four clean-up sites on Table 6, sites shown on Figure 1.

Interval Calculation	All San Diego Bay (n = 161)		Clean-up Areas (n = 87)	
	Lower (SQO score > 2)	**Upper (SQO score < 3)**	**Lower (SQO score > 2)**	**Upper (SQO score < 3)**
Confidence Interval: 95%	0.280	0.250	0.398	0.295
Confidence Interval: 90%	0.293	0.240	0.418	0.273
Confidence Interval: 80%	0.309	0.229	0.443	0.249
Prediction Interval: 95% (n = 3)	0.113	0.511	0.189	0.654
Prediction Interval: 90% (n = 3)	0.145	0.418	0.234	0.504
Prediction Interval: 80% (n = 3)	0.195	0.329	0.303	0.371
Prediction Interval: 95% (n = 5)	0.144	0.421	0.231	0.519
Prediction Interval: 90% (n = 5)	0.174	0.360	0.274	0.422
Prediction Interval: 80% (n = 5)	0.220	0.298	0.336	0.331

Table 9. Impact limits for copper. Site-specific limits used samples from Southwest and Paleta Creek as identified on Table 6 and shown on Figure 1.

Interval Calculation	All San Diego Bay (n = 161)		Clean-up Areas (n = 87)		Site-specific (n = 49)	
	Lower (SQO score > 2)	**Upper (SQO score < 3)**	**Lower (SQO score > 2)**	**Upper (SQO score < 3)**	**Lower (SQO score > 2)**	**Upper (SQO score < 3)**
Confidence Interval: 95%	138.440	90.878	137.799	151.363	135.212	159.049
Confidence Interval: 90%	141.598	87.646	141.774	140.580	142.749	146.349
Confidence Interval: 80%	145.485	83.925	146.673	128.845	152.137	132.779
Prediction Interval: 95% (n = 3)	87.525	173.588	88.781	320.555	82.113	320.800
Prediction Interval: 90% (n = 3)	99.171	144.840	100.822	250.869	97.196	251.097
Prediction Interval: 80% (n = 3)	115.216	116.596	117.365	187.887	118.484	188.550
Prediction Interval: 95% (n = 5)	98.826	145.859	100.138	257.947	95.060	260.445
Prediction Interval: 90% (n = 5)	108.981	126.537	110.686	212.133	108.806	213.891
Prediction Interval: 80% (n = 5)	122.558	106.729	124.752	168.443	127.506	169.896

Table 10. Impact limits for lead. Site-specific limits used samples from Southwest, NASSCO, and Paleta Creek as identified on Table 6, and shown on Figure 1.

Interval Calculation	All San Diego Bay (n = 161)		Clean-up Areas (n = 87)		Site-specific (n = 63)	
	Lower (SQO score > 2)	**Upper (SQO score < 3)**	**Lower (SQO score > 2)**	**Upper (SQO score < 3)**	**Lower (SQO score > 2)**	**Upper (SQO score < 3)**
Confidence Interval: 95%	56.753	35.516	68.926	58.652	65.059	59.847
Confidence Interval: 90%	58.032	34.333	70.644	55.217	67.497	56.158
Confidence Interval: 80%	59.605	32.968	72.751	51.423	70.507	52.116
Prediction Interval: 95% (n = 3)	36.082	65.062	47.118	108.244	42.309	107.418
Prediction Interval: 90% (n = 3)	40.821	54.925	52.599	88.608	48.393	88.142
Prediction Interval: 80% (n = 3)	47.339	44.839	59.989	69.975	56.749	69.873
Prediction Interval: 95% (n = 5)	40.681	55.287	52.291	90.644	47.810	90.470
Prediction Interval: 90% (n = 5)	44.807	48.406	57.024	77.266	53.190	77.217
Prediction Interval: 80% (n = 5)	50.318	41.280	63.243	64.003	60.358	64.111

Table 11. Impact limits for mercury. Site-specific limits used samples from Southwest and Paleta Creek as identified on Table 6, and shown on Figure 1.

Interval Calculation	All San Diego Bay (n = 161)		Clean-up Areas (n = 87)		Site-specific (n = 49)	
	Lower (SQO score > 2)	**Upper (SQO score < 3)**	**Lower (SQO score > 2)**	**Upper (SQO score < 3)**	**Lower (SQO score > 2)**	**Upper (SQO score < 3)**
Confidence Interval: 95%	0.433	0.366	0.406	0.552	0.465	0.603
Confidence Interval: 90%	0.445	0.354	0.419	0.518	0.482	0.562
Confidence Interval: 80%	0.460	0.339	0.435	0.480	0.502	0.518
Prediction Interval: 95% (n = 3)	0.251	0.680	0.249	1.059	0.336	1.095
Prediction Interval: 90% (n = 3)	0.291	0.572	0.287	0.856	0.375	0.889
Prediction Interval: 80% (n = 3)	0.348	0.465	0.340	0.666	0.427	0.697
Prediction Interval: 95% (n = 5)	0.290	0.576	0.285	0.877	0.370	0.917
Prediction Interval: 90% (n = 5)	0.326	0.502	0.318	0.740	0.404	0.776
Prediction Interval: 80% (n = 5)	0.375	0.427	0.363	0.606	0.448	0.638

Table 12. Impact limits for zinc. Site-specific limits used samples from Southwest, Paleta Creek, and Chollas Creek as identified on Table 6, and shown on Figure 1.

Interval Calculation	All San Diego Bay (n = 161)		Clean-up Areas (n = 87)		Site-specific (n = 73)	
	Lower (SQO score > 2)	Upper (SQO score < 3)	Lower (SQO score > 2)	Upper (SQO score < 3)	Lower (SQO score > 2)	Upper (SQO score < 3)
Confidence Interval: 95%	234.723	153.627	255.105	225.592	252.783	214.277
Confidence Interval: 90%	239.407	149.586	262.050	214.160	261.227	203.230
Confidence Interval: 80%	245.155	144.886	270.595	201.419	271.649	190.968
Prediction Interval: 95% (n = 3)	157.076	247.359	168.416	382.506	162.297	353.177
Prediction Interval: 90% (n = 3)	175.241	216.498	189.913	321.906	185.346	298.745
Prediction Interval: 80% (n = 3)	199.844	184.553	219.219	262.653	217.094	245.376
Prediction Interval: 95% (n = 5)	174.708	217.619	188.697	328.270	183.631	305.152
Prediction Interval: 90% (n = 5)	190.337	196.008	207.417	286.072	203.949	266.909
Prediction Interval: 80% (n = 5)	210.958	172.925	232.230	243.219	231.091	228.028

Table 13. Impact limits for chlordanes. Impact limits were based only on samples from Chollas Creek and Paleta Creek; chlordane was not sampled at Southwest Marine or NASSCO.

Interval Calculation	All San Diego Bay (n = 161)		Clean-up Areas (n = 57)	
	Lower (SQO score > 2)	Upper (SQO score < 3)	Lower (SQO score > 2)	Upper (SQO score < 3)
Confidence Interval: 95%	3.651	1.811	8.939	2.039
Confidence Interval: 90%	3.916	1.707	9.630	1.795
Confidence Interval: 80%	4.260	1.591	10.521	1.549
Prediction Interval: 95% (n = 3)	1.053	4.508	3.624	5.009
Prediction Interval: 90% (n = 3)	1.492	3.463	4.790	3.574
Prediction Interval: 80% (n = 3)	2.265	2.527	6.670	2.419
Prediction Interval: 95% (n = 5)	1.472	3.512	4.678	3.810
Prediction Interval: 90% (n = 5)	1.935	2.853	5.836	2.898
Prediction Interval: 80% (n = 5)	2.685	2.226	7.588	2.112

Table 14. Impact limits for DDTs. Impact limits were based only on samples from Chollas Creek and Paleta Creek; DDTs were not measured at Southwest Marine or NASSCO.

Interval Calculation	All San Diego Bay (n = 161)		Clean-up Areas (n = 58)	
	Lower (SQO score > 2)	**Upper (SQO score < 3)**	**Lower (SQO score > 2)**	**Upper (SQO score < 3)**
Confidence Interval: 95%	5.510	3.056	12.196	4.134
Confidence Interval: 90%	5.868	2.884	13.043	3.698
Confidence Interval: 80%	6.328	2.692	14.125	3.250
Prediction Interval: 95% (n = 3)	1.786	7.473	5.336	9.078
Prediction Interval: 90% (n = 3)	2.448	5.769	6.881	6.757
Prediction Interval: 80% (n = 3)	3.571	4.236	9.304	4.801
Prediction Interval: 95% (n = 5)	2.418	5.850	6.738	7.145
Prediction Interval: 90% (n = 5)	3.097	4.772	8.241	5.624
Prediction Interval: 80% (n = 5)	4.165	3.741	10.467	4.264

Table 15. Impact limits for LPAHs. Site-specific limits were not calculated because LPAHs was indicated at all four clean-up sites on Table 6, sites shown on Figure 1.

Interval Calculation	All San Diego Bay (n = 161)		Clean-up Areas (n = 87)	
	Lower (SQO score > 2)	**Upper (SQO score < 3)**	**Lower (SQO score > 2)**	**Upper (SQO score < 3)**
Confidence Interval: 95%	158.504	130.904	286.634	204.793
Confidence Interval: 90%	167.502	122.975	303.026	181.588
Confidence Interval: 80%	178.980	114.106	323.841	157.576
Prediction Interval: 95% (n = 3)	51.601	399.880	121.333	694.360
Prediction Interval: 90% (n = 3)	70.055	292.582	155.593	465.976
Prediction Interval: 80% (n = 3)	101.123	201.233	209.419	291.118
Prediction Interval: 95% (n = 5)	69.462	296.146	153.537	487.554
Prediction Interval: 90% (n = 5)	88.249	231.747	186.750	354.683
Prediction Interval: 80% (n = 5)	117.631	172.757	235.970	243.708

Table 16. Impact limits for HPAHs. Site-specific limits were not calculated because HPAHs were indicated at all four clean-up sites on Table 6, sites shown on Figure 1.

Interval Calculation	All San Diego Bay (n = 161)		Clean-up Areas (n = 87)	
	Lower (SQO score > 2)	**Upper (SQO score < 3)**	**Lower (SQO score > 2)**	**Upper (SQO score < 3)**
Confidence Interval: 95%	922.138	572.579	2369.871	1349.352
Confidence Interval: 90%	984.249	530.502	2498.070	1172.617
Confidence Interval: 80%	1064.365	484.144	2660.334	993.689
Prediction Interval: 95% (n = 3)	245.114	2240.066	1049.641	5612.283
Prediction Interval: 90% (n = 3)	351.672	1529.391	1328.503	3523.132
Prediction Interval: 80% (n = 3)	542.427	968.178	1760.327	2034.435
Prediction Interval: 95% (n = 5)	348.156	1552.174	1311.867	3714.314
Prediction Interval: 90% (n = 5)	461.870	1150.420	1579.276	2561.948
Prediction Interval: 80% (n = 5)	648.452	803.549	1971.073	1653.198

Table 17. Impact limits for PCBs. Site-specific limits used samples from Southwest, NASSCO, and Paleta Creek as identified on Table 6, and shown on Figure 1.

Interval Calculation	All San Diego Bay (n = 161)		Clean-up Areas (n = 87)		Site-specific (n = 63)	
	Lower (SQO score > 2)	**Upper (SQO score < 3)**	**Lower (SQO score > 2)**	**Upper (SQO score < 3)**	**Lower (SQO score > 2)**	**Upper (SQO score < 3)**
Confidence Interval: 95%	42.843	27.839	135.532	116.661	141.093	150.907
Confidence Interval: 90%	45.960	25.772	142.329	101.156	151.299	128.544
Confidence Interval: 80%	50.002	23.496	150.896	85.496	164.358	106.476
Prediction Interval: 95% (n = 3)	10.281	110.567	63.614	496.313	62.347	659.525
Prediction Interval: 90% (n = 3)	15.167	75.172	79.175	309.271	80.457	400.551
Prediction Interval: 80% (n = 3)	24.190	47.347	102.828	177.041	108.857	223.009
Prediction Interval: 95% (n = 5)	15.004	76.304	78.254	326.327	78.628	427.762
Prediction Interval: 90% (n = 5)	20.343	56.367	92.968	223.762	96.267	286.918
Prediction Interval: 80% (n = 5)	29.320	39.216	114.215	143.392	122.373	179.503

Table 18. Selected sediment guidelines and thresholds from other studies for chemicals evaluated in San Diego Bay. See footnotes for key to superscripts.

Chemical	Units	ERL[1]	ERM[1]	Estimated Benthic Impact Limit*	LC50s	LOELs[5]	Upper Prediction Limits
Cadmium	ug/g	1.2	9.6	--	> 1000 (2)	--	--
Chlordane	ng/g	0.5	6	32		--	0.3 (6)
Copper	ug/g	34	270	--	60.70 (2) 534.3 (3)	1000	1900 (5)
DDT	ng/g	1.58	46.1	25	55 (4)	--	21 (6)
HPAH	ng/g	1700	9600	4900	--	27000	26000 (5)
Lead	ug/g	46.7	218	83	--	250	480 (5)
LPAH	ng/g	552	3160	500	--	--	--
Mercury	ug/g	0.15	0.71	--	--	3.2	2.5 (5)
PCB	ng/g	22.7	180	--	--	3000	5800 (5) 84 (6)
Zinc	ug/g	150	410	330	--	1200	4600 (5)

* See Methods

Sources: [1] Long *et al.* (1995); [2] Weston (1996); [3] Anderson *et al.* (2008);
[4] Swartz *et al.* (1994); [5] Exponent (2002); [6] SCCWRP (2005)

Table 19. Impact limits for mERMq.

Interval Calculation	All San Diego Bay (n = 161)		Clean-up Areas (n = 87)	
	Lower (SQO score > 2)	**Upper (SQO score < 3)**	**Lower (SQO score > 2)**	**Upper (SQO score < 3)**
Confidence Interval: 95%	0.292	0.191	0.381	0.277
Confidence Interval: 90%	0.301	0.185	0.394	0.260
Confidence Interval: 80%	0.311	0.178	0.411	0.242
Prediction Interval: 95% (n = 3)	0.164	0.339	0.227	0.523
Prediction Interval: 90% (n = 3)	0.192	0.289	0.263	0.425
Prediction Interval: 80% (n = 3)	0.232	0.238	0.315	0.333
Prediction Interval: 95% (n = 5)	0.191	0.291	0.261	0.435
Prediction Interval: 90% (n = 5)	0.216	0.256	0.294	0.369
Prediction Interval: 80% (n = 5)	0.251	0.220	0.339	0.304

Table 20. Impact limits for mCSI.

Interval Calculation	All San Diego Bay (n = 161)		Clean-up Areas (n = 87)	
	Lower (SQO score > 2)	**Upper (SQO score < 3)**	**Lower (SQO score > 2)**	**Upper (SQO score < 3)**
Confidence Interval: 95%	2.509	1.972	2.748	2.399
Confidence Interval: 90%	2.531	1.944	2.773	2.334
Confidence Interval: 80%	2.557	1.910	2.802	2.258
Prediction Interval: 95% (n = 3)	2.108	2.561	2.399	3.180
Prediction Interval: 90% (n = 3)	2.210	2.380	2.495	2.900
Prediction Interval: 80% (n = 3)	2.340	2.181	2.615	2.602
Prediction Interval: 95% (n = 5)	2.207	2.387	2.490	2.931
Prediction Interval: 90% (n = 5)	2.291	2.254	2.568	2.723
Prediction Interval: 80% (n = 5)	2.396	2.104	2.665	2.497

Table 21. Impact limits for CA LRM.

Interval Calculation	All San Diego Bay (n = 161)		Clean-up Areas (n = 87)	
	Lower (SQO score > 2)	Upper (SQO score < 3)	Lower (SQO score > 2)	Upper (SQO score < 3)
Confidence Interval: 95%	0.661	0.524	0.690	0.609
Confidence Interval: 90%	0.665	0.517	0.696	0.595
Confidence Interval: 80%	0.671	0.509	0.703	0.578
Prediction Interval: 95% (n = 3)	0.575	0.661	0.606	0.772
Prediction Interval: 90% (n = 3)	0.598	0.620	0.629	0.714
Prediction Interval: 80% (n = 3)	0.625	0.573	0.658	0.652
Prediction Interval: 95% (n = 5)	0.597	0.621	0.628	0.720
Prediction Interval: 90% (n = 5)	0.615	0.590	0.650	0.677
Prediction Interval: 80% (n = 5)	0.637	0.555	0.670	0.630

FIGURES

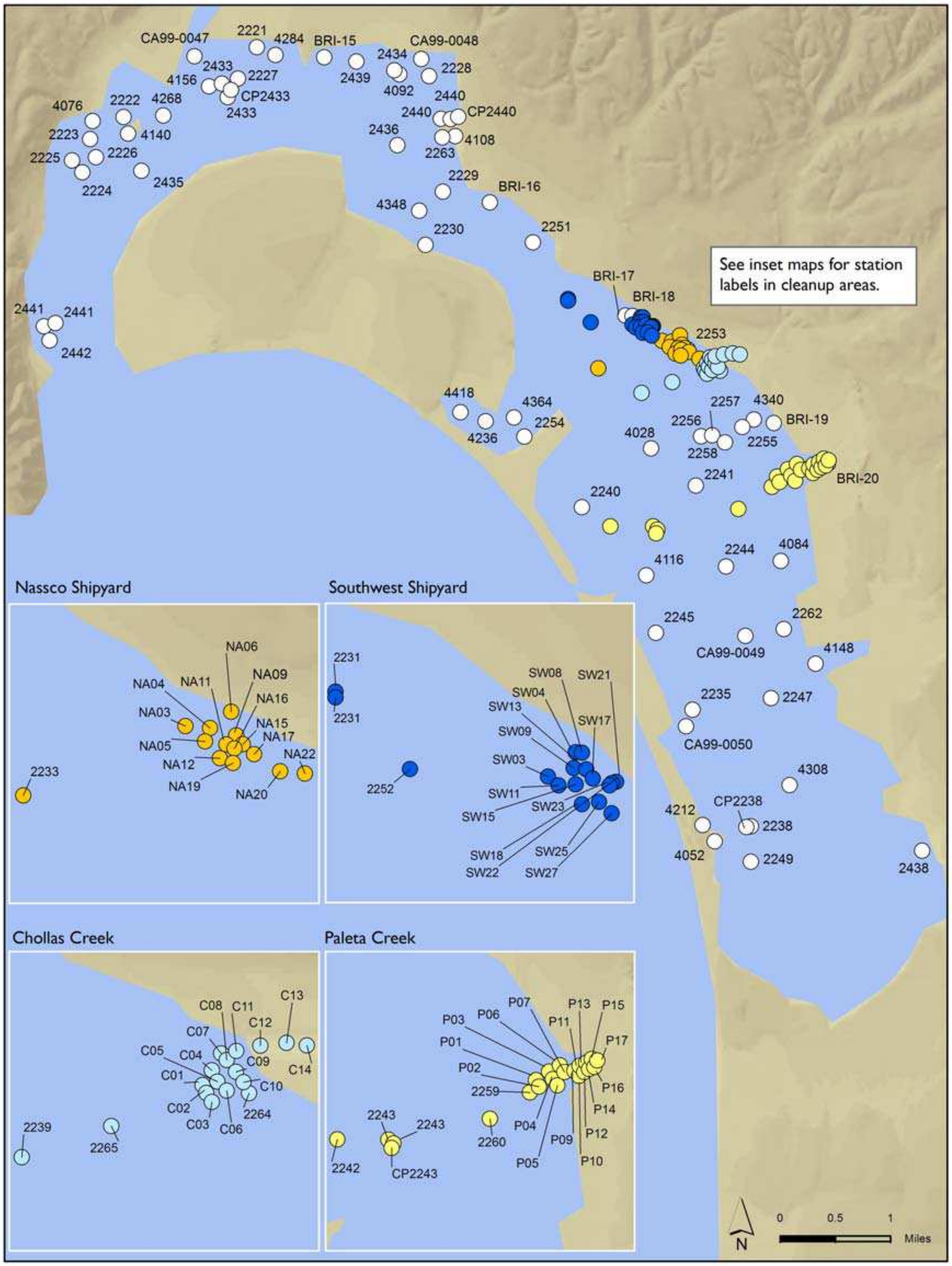

Figure 1. Map of San Diego Bay Stations. All stations were used in analyses conducted at the San Diego Bay spatial scale. Clean-up area stations are identified by colors on main map and insets. Note that in some areas (*e.g.,* Station 2243 in Paleta Creek) duplicate stations exist because the same station was sampled at different times in another study.

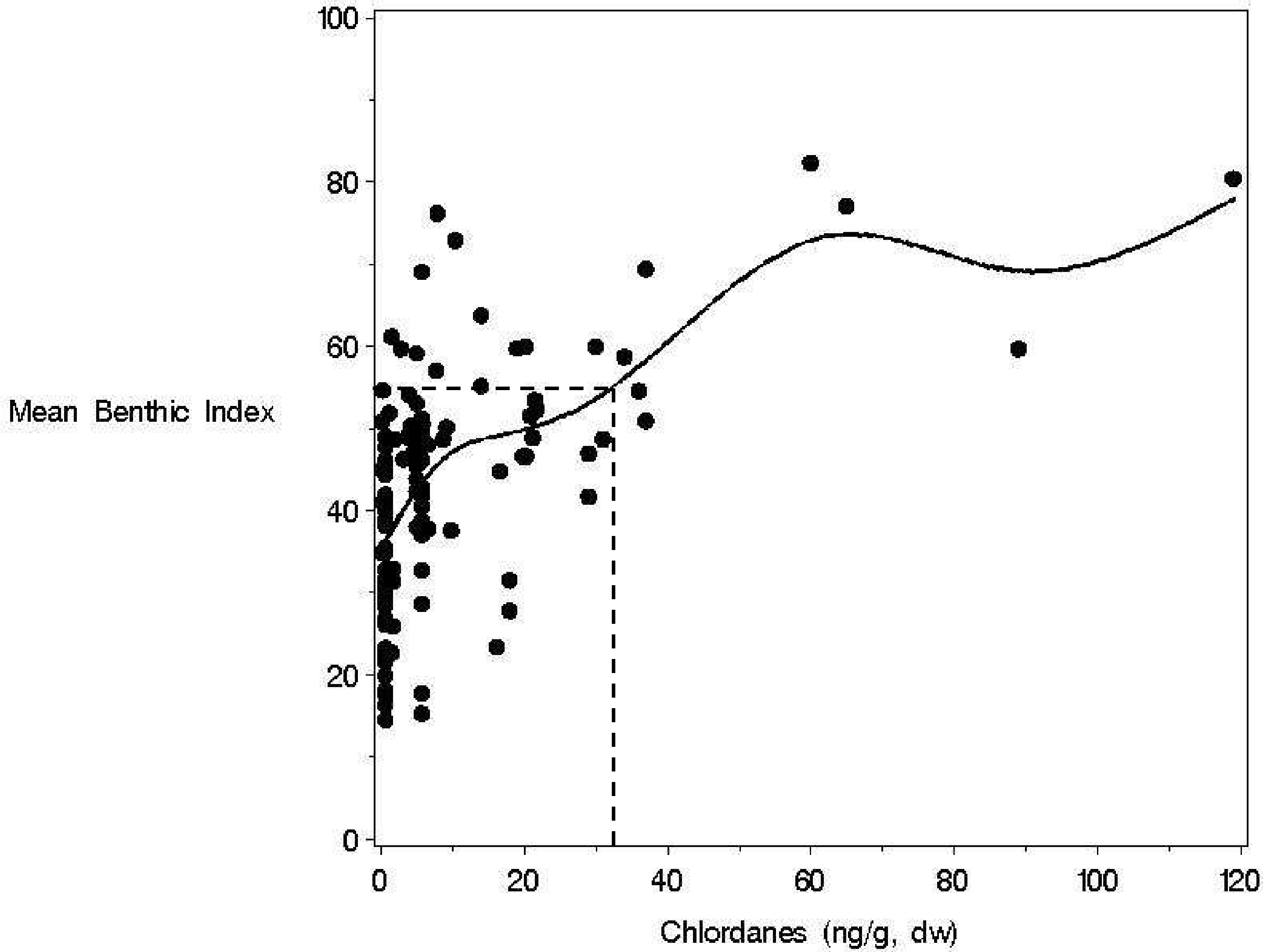

Figure 2. Example of the graphical method used to estimate benthic impact limits. An empirical limit of 32 ng/g was determined by finding the concentration value on the fitted curve that corresponded to a mean benthic index value of 55 (the index value above which all samples were 'impacted').

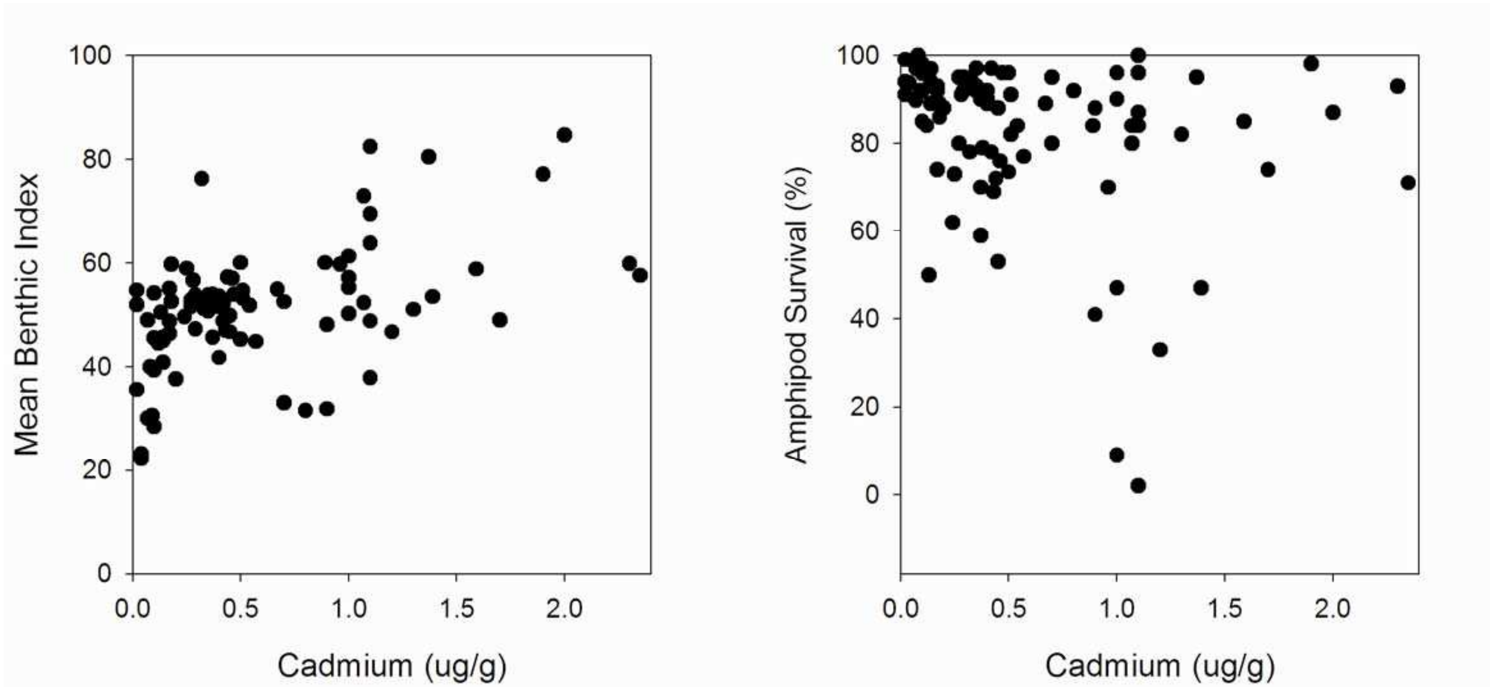

Figure 3. Relationships between cadmium : mean benthic index, and cadmium : control-adjusted percent amphipod survival in sediments from clean-up areas (n = 87).

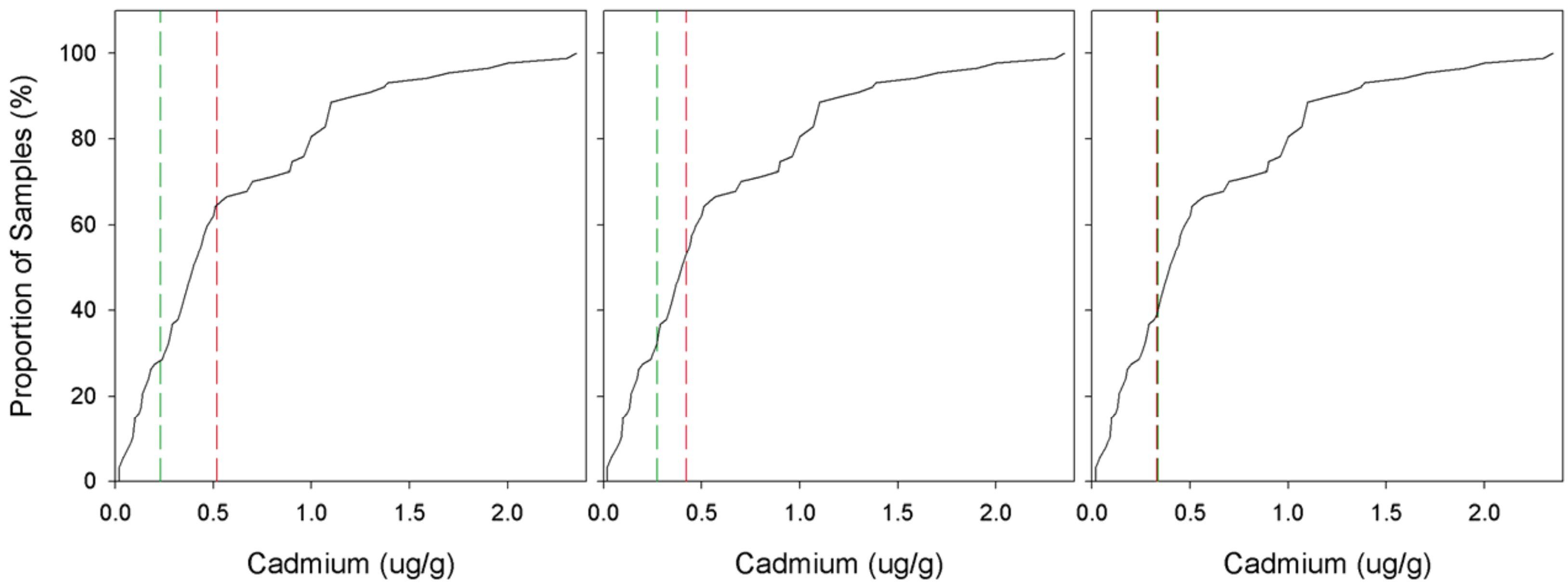

Figure 4. Cumulative distribution function (CDF) for cadmium in sediments from clean-up areas (n = 87). Dashed lines indicate prediction limits for 5 future samples at probability levels of 95% (left plot), 90% (center plot), and 80% (right plot). Red lines are upper limits calculated from un-impacted samples (SQO scores = 1 – 2). Green lines are lower limits calculated from impacted samples (SQO scores = 3 – 5). The proportion of future samples below the red line are representative of the population of un-impacted samples with 95%, 90%, or 80% statistical confidence, whereas the proportion of future samples above the green line are representative of the population of impacted samples at that statistical level.

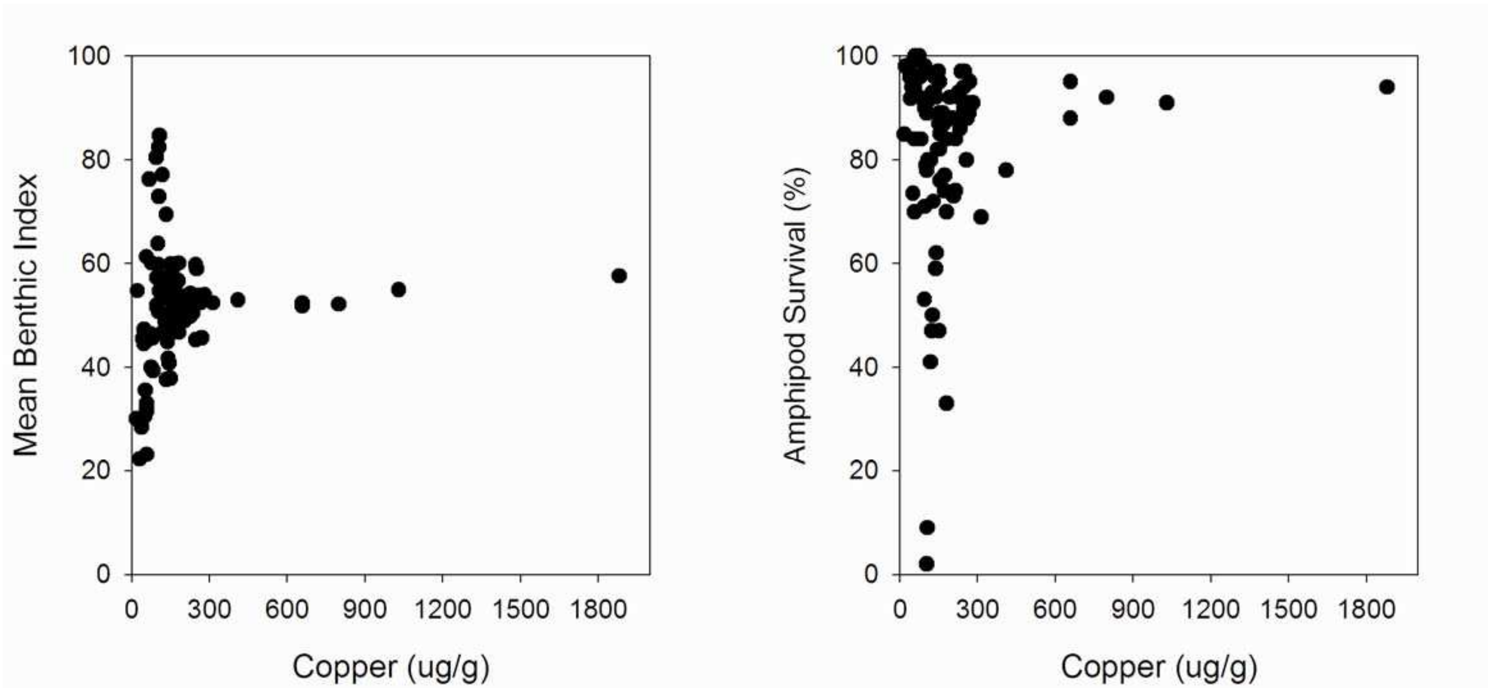

Figure 5. Relationships between copper : mean benthic index, and copper : control-adjusted percent amphipod survival in sediments from clean-up areas (n = 87).

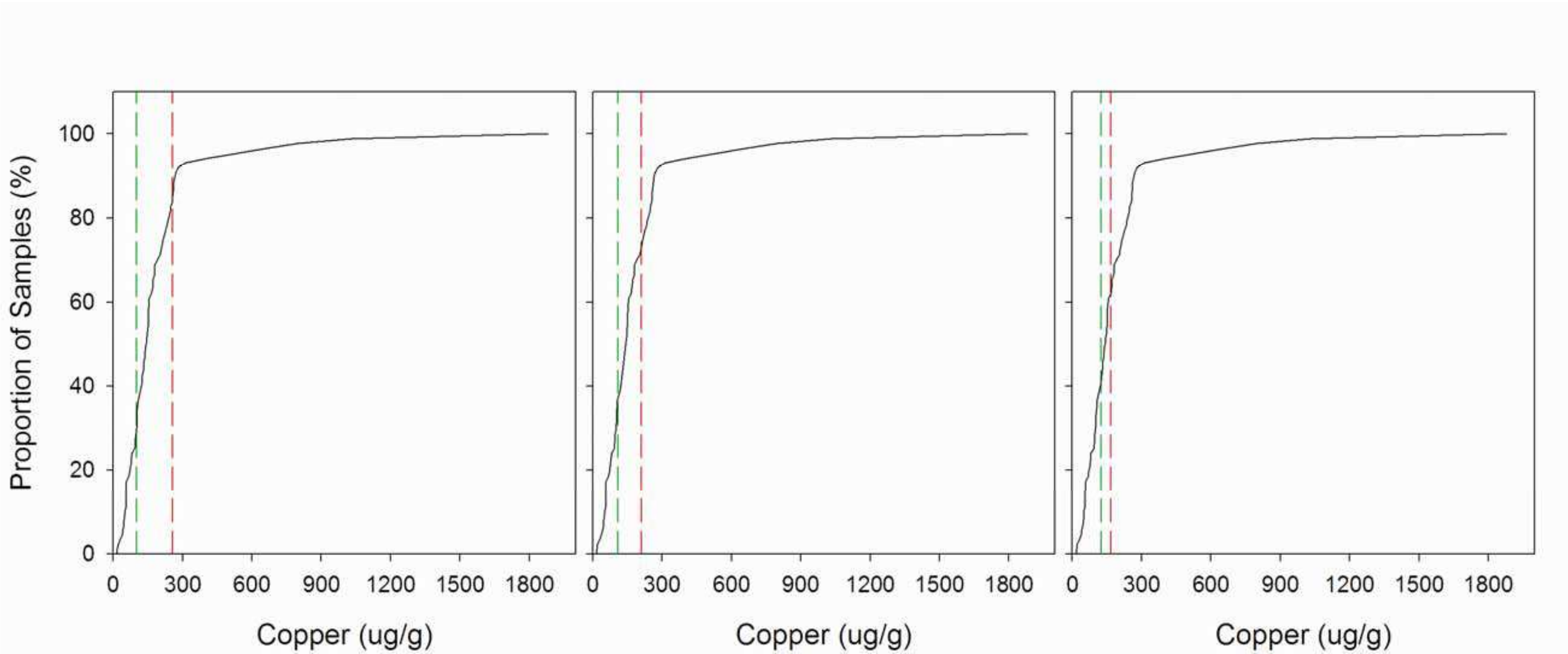

Figure 6. Cumulative distribution function (CDF) for copper in sediments from clean-up areas (n = 87). Dashed lines indicate prediction limits for 5 future samples at probability levels of 95% (left plot), 90% (center plot), and 80% (right plot). Red lines are upper limits calculated from un-impacted samples (SQO scores = 1 – 2). Green lines are lower limits calculated from impacted samples (SQO scores = 3 – 5). The proportion of future samples below the red line are representative of the population of un-impacted samples with 95%, 90%, or 80% statistical confidence, whereas the proportion of future samples above the green line are representative of the population of impacted samples at that statistical level.

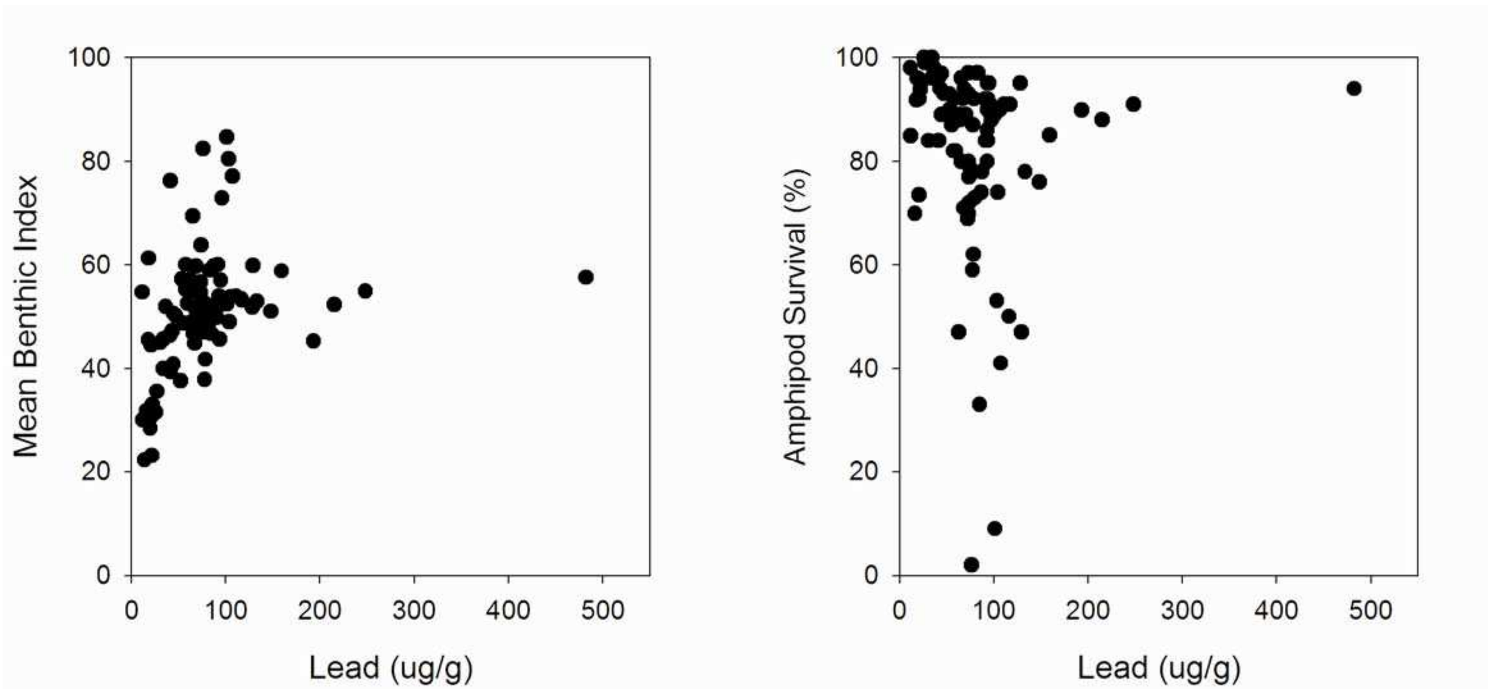

Figure 7. Relationships between lead : mean benthic index, and lead : control-adjusted percent amphipod survival in sediments from clean-up areas (n = 87).

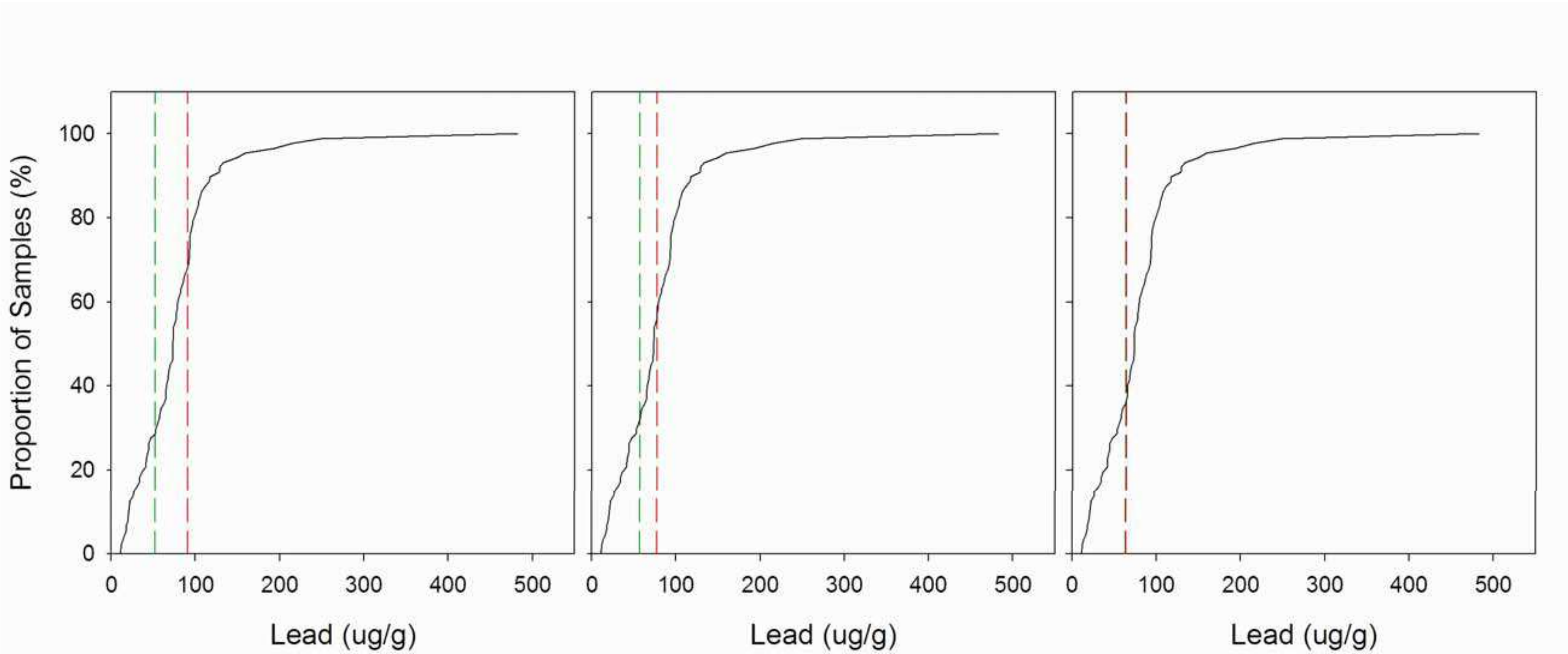

Figure 8. Cumulative distribution function (CDF) for lead in sediments from clean-up areas (n = 87). Dashed lines indicate prediction limits for 5 future samples at probability levels of 95% (left plot), 90% (center plot), and 80% (right plot). Red lines are upper limits calculated from un-impacted samples (SQO scores = 1 – 2). Green lines are lower limits calculated from impacted samples (SQO scores = 3 – 5). The proportion of future samples below the red line are representative of the population of un-impacted samples with 95%, 90%, or 80% statistical confidence, whereas the proportion of future samples above the green line are representative of the population of impacted samples at that statistical level.

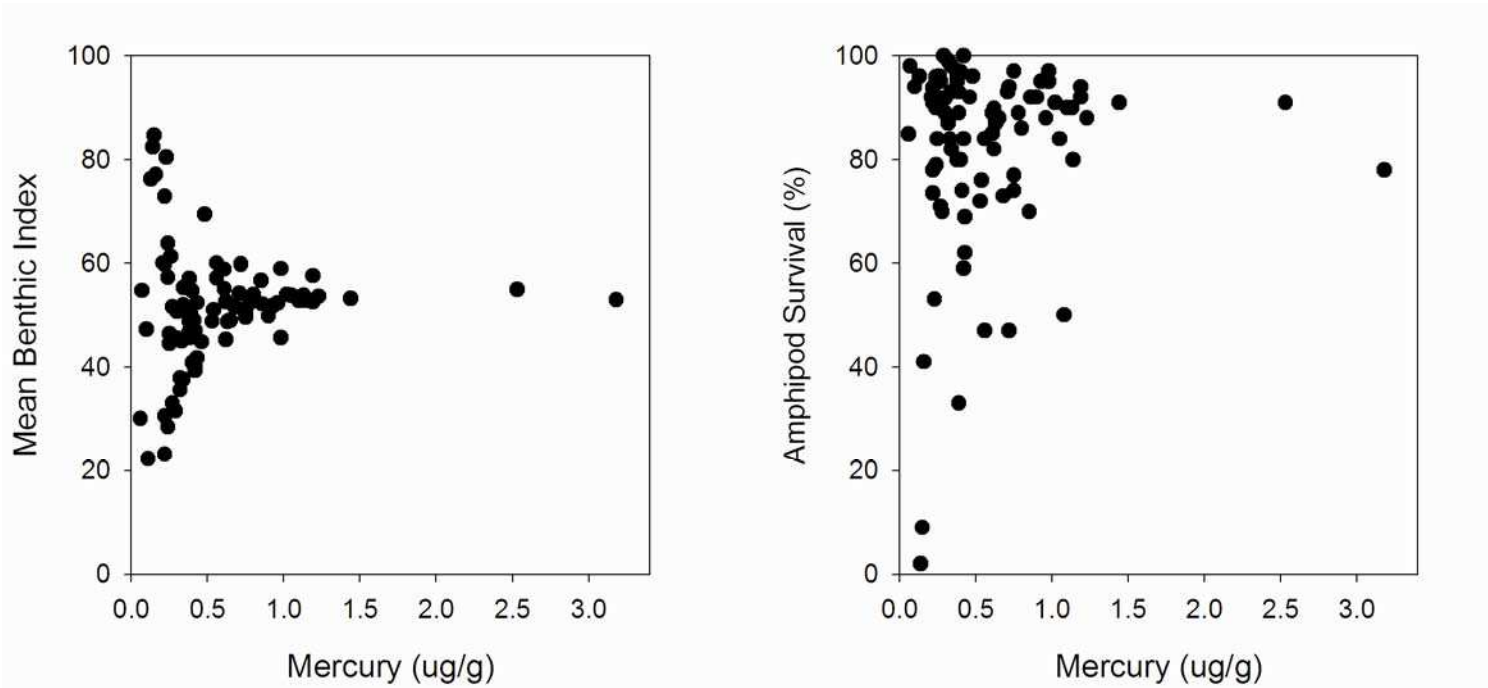

Figure 9. Relationships between mercury : mean benthic index, and mercury : control-adjusted percent amphipod survival in sediments from clean-up areas (n = 87).

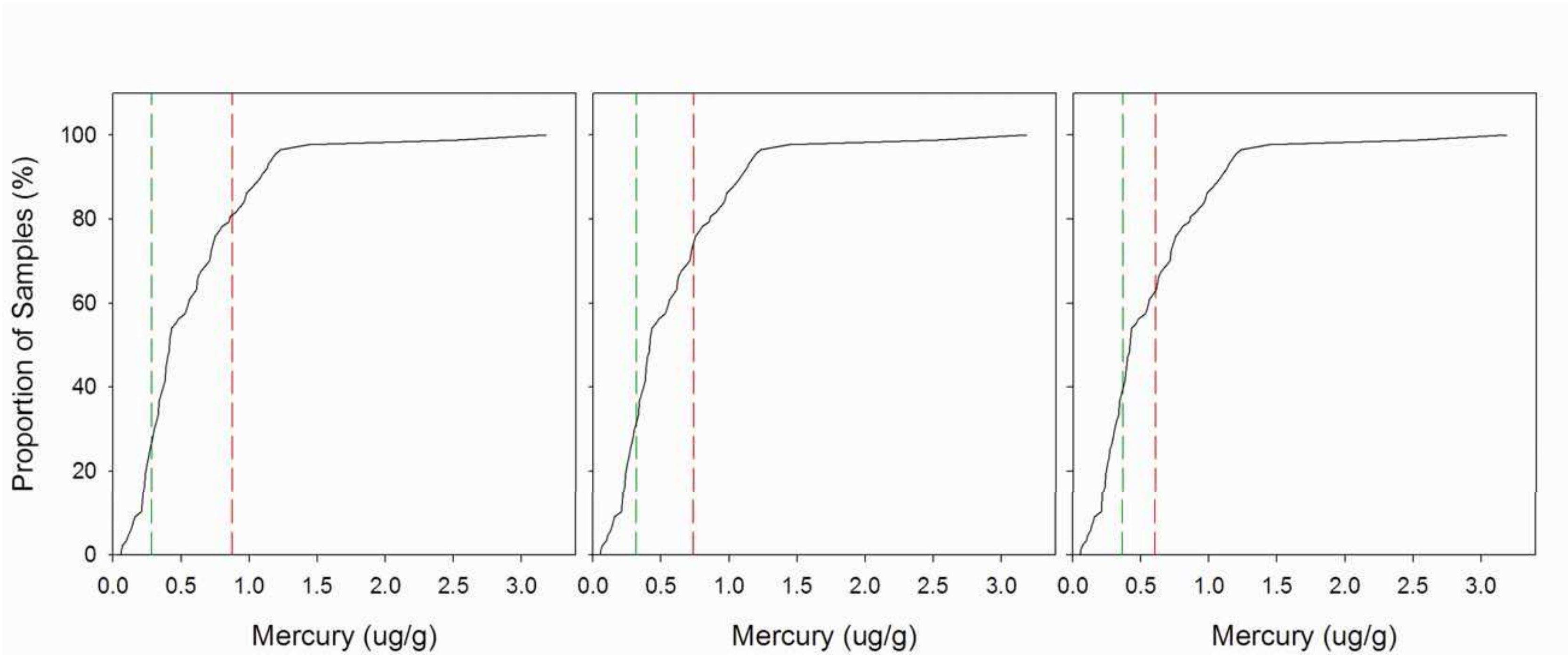

Figure 10. Cumulative distribution function (CDF) for mercury in sediments from clean-up areas (n = 87). Dashed lines indicate prediction limits for 5 future samples at probability levels of 95% (left plot), 90% (center plot), and 80% (right plot). Red lines are upper limits calculated from un-impacted samples (SQO scores = 1 – 2). Green lines are lower limits calculated from impacted samples (SQO scores = 3 – 5). The proportion of future samples below the red line are representative of the population of un-impacted samples with 95%, 90%, or 80% statistical confidence, whereas the proportion of future samples above the green line are representative of the population of impacted samples at that statistical level.

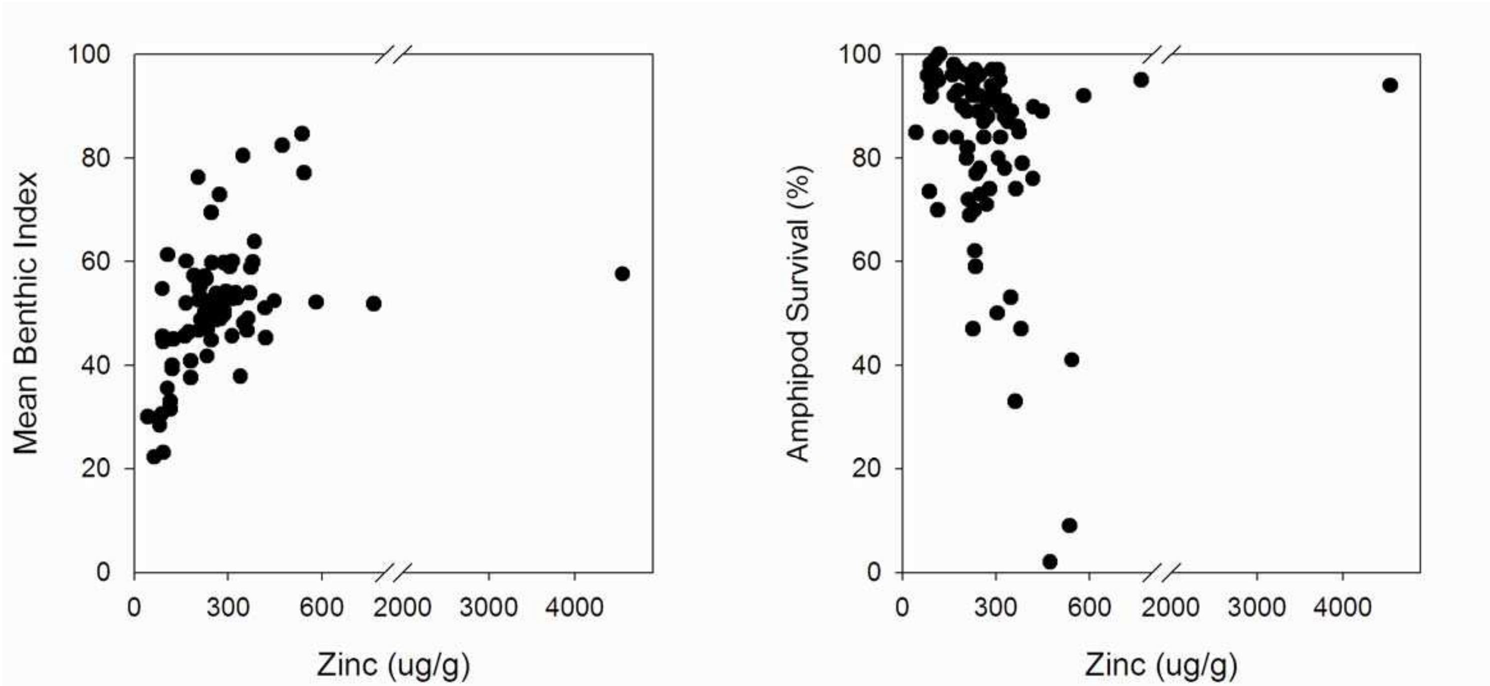

Figure 11. Relationships between zinc: mean benthic index, and zinc : control-adjusted percent amphipod survival in sediments from clean-up areas (n = 87).

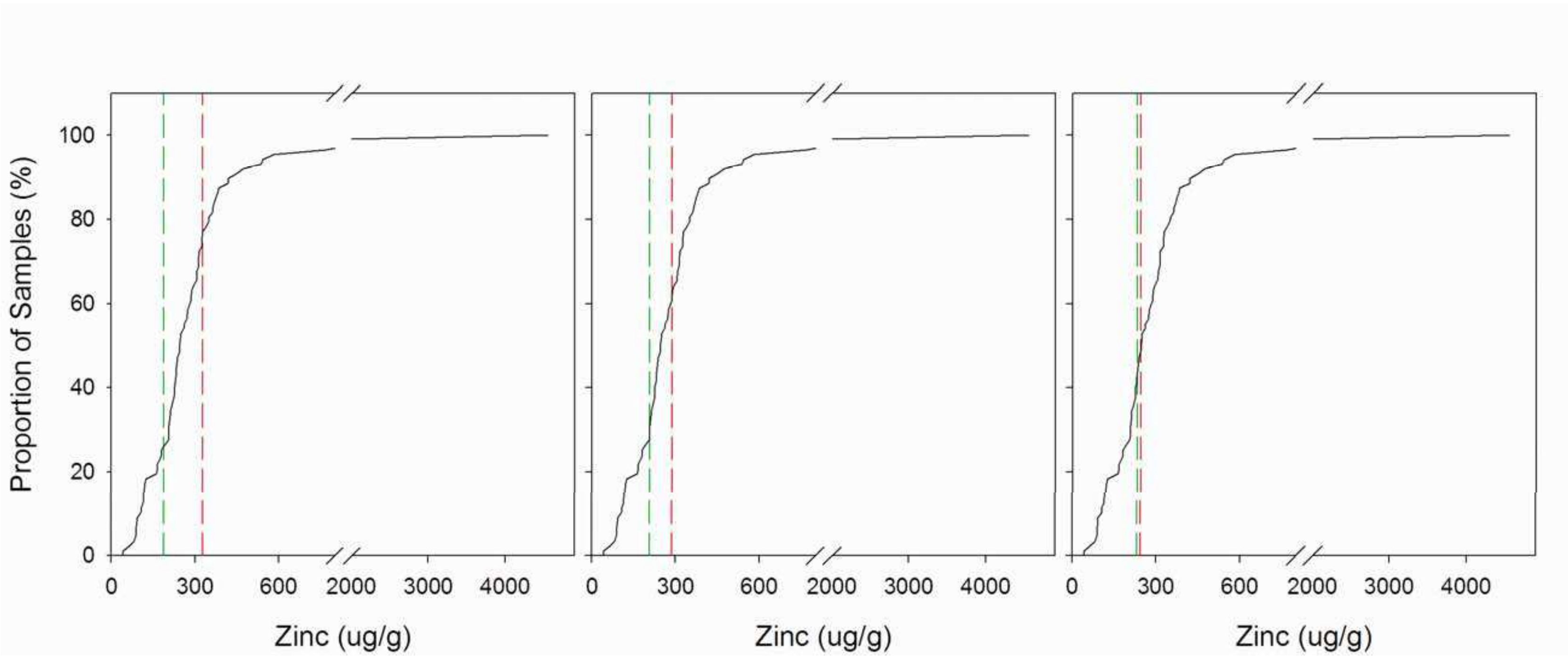

Figure 12. Cumulative distribution function (CDF) for zinc in sediments from clean-up areas (n = 87). Dashed lines indicate prediction limits for 5 future samples at probability levels of 95% (left plot), 90% (center plot), and 80% (right plot). Red lines are upper limits calculated from un-impacted samples (SQO scores = 1 – 2). Green lines are lower limits calculated from impacted samples (SQO scores = 3 – 5). The proportion of future samples below the red line are representative of the population of un-impacted samples with 95%, 90%, or 80% statistical confidence, whereas the proportion of future samples above the green line are representative of the population of impacted samples at that statistical level.

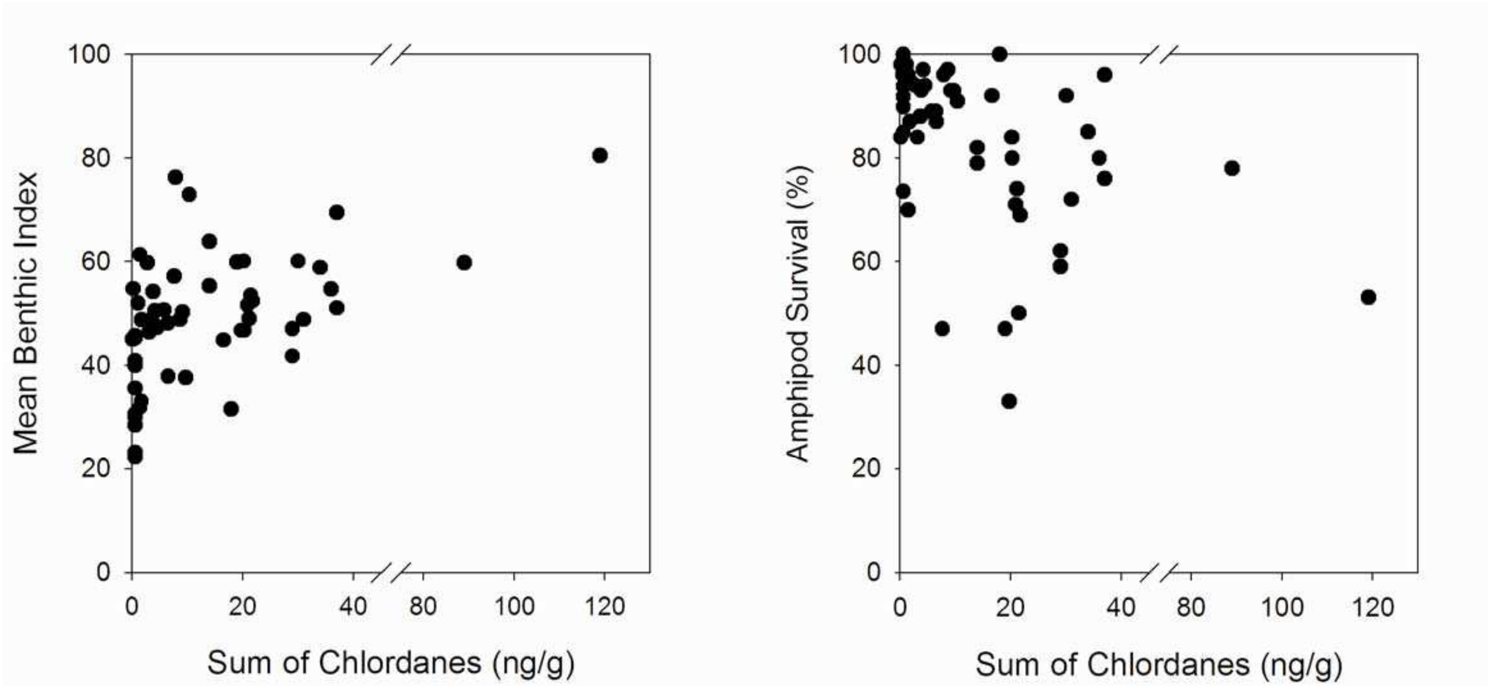

Figure 13. Relationships between chlordanes : mean benthic index, and chlordanes : control-adjusted percent amphipod survival in sediments from clean-up areas (n = 57).

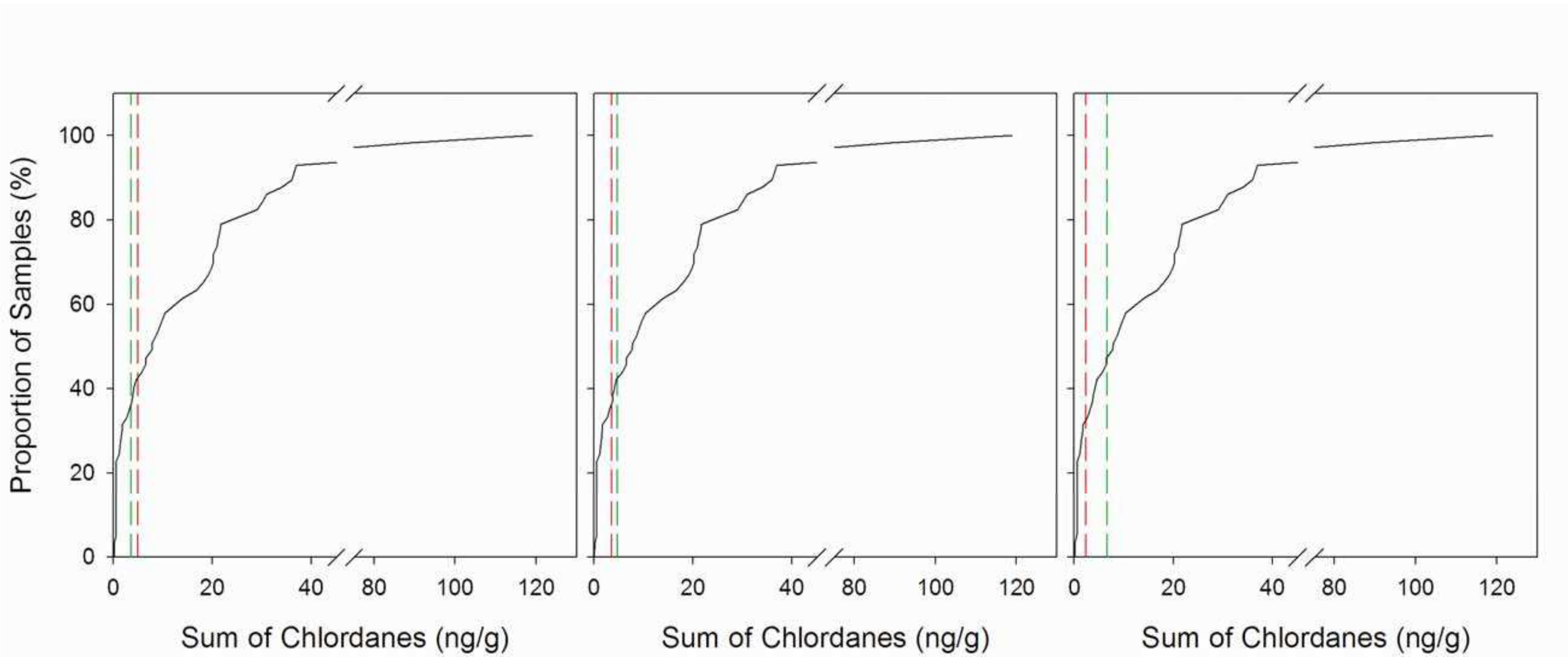

Figure 14. Cumulative distribution function (CDF) for chlordanes in sediments from clean-up areas (n = 57). Dashed lines indicate prediction limits for 5 future samples at probability levels of 95% (left plot), 90% (center plot), and 80% (right plot). Red lines are upper limits calculated from un-impacted samples (SQO scores = 1 – 2). Green lines are lower limits calculated from impacted samples (SQO scores = 3 – 5). The proportion of future samples below the red line are representative of the population of un-impacted samples with 95%, 90%, or 80% statistical confidence, whereas the proportion of future samples above the green line are representative of the population of impacted samples at that statistical level.

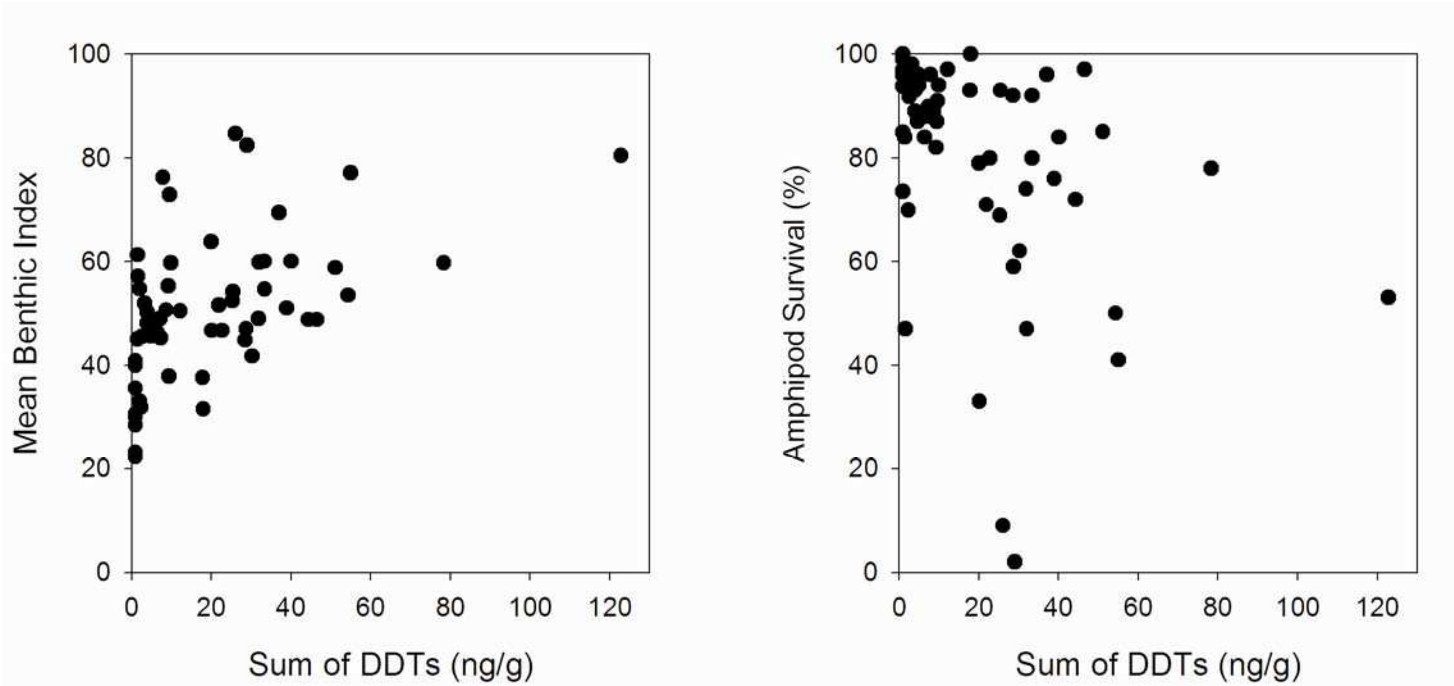

Figure 15. Relationships between DDTs : mean benthic index, and DDTs : control-adjusted percent amphipod survival in sediments from clean-up areas (n = 58).

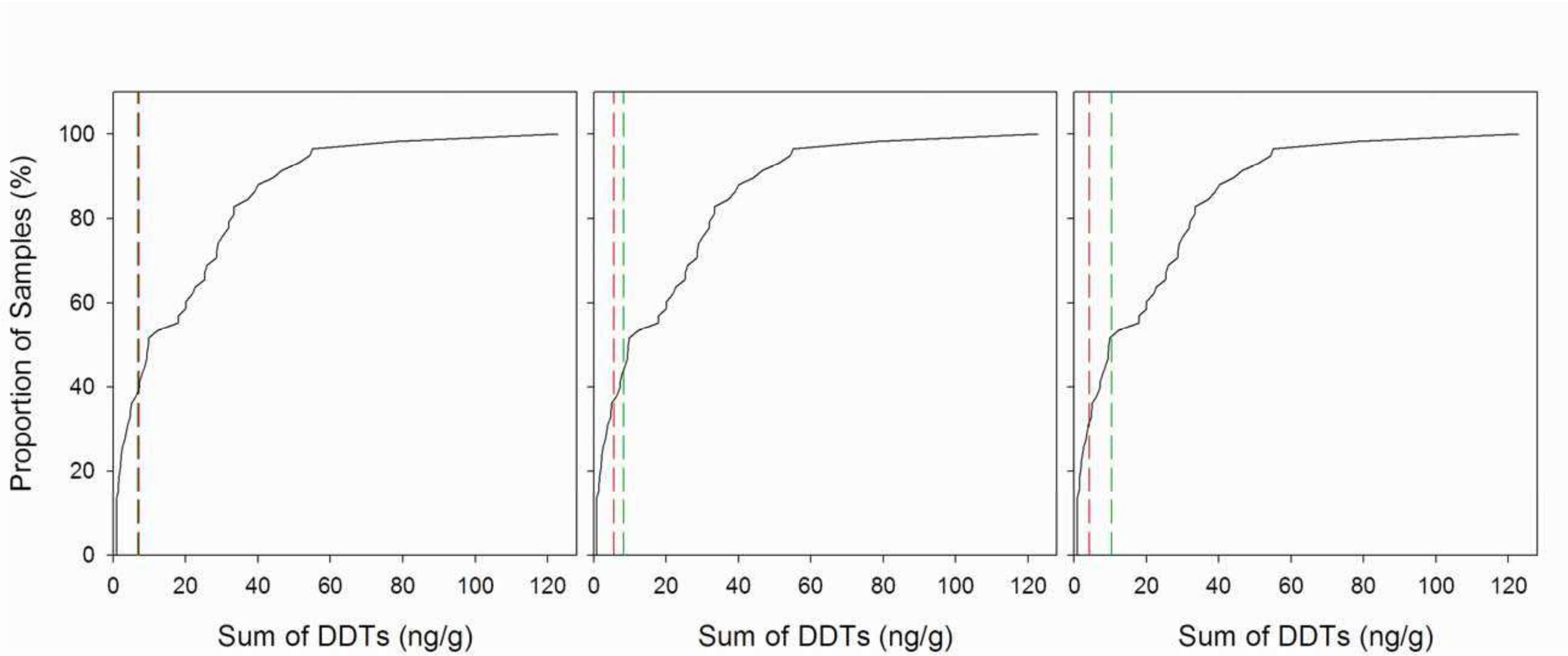

Figure 16. Cumulative distribution function (CDF) for DDTs in sediments from clean-up areas (n = 58). Dashed lines indicate prediction limits for 5 future samples at probability levels of 95% (left plot), 90% (center plot), and 80% (right plot). Red lines are upper limits calculated from un-impacted samples (SQO scores = 1 – 2). Green lines are lower limits calculated from impacted samples (SQO scores = 3 – 5). The proportion of future samples below the red line are representative of the population of un-impacted samples with 95%, 90%, or 80% statistical confidence, whereas the proportion of future samples above the green line are representative of the population of impacted samples at that statistical level.

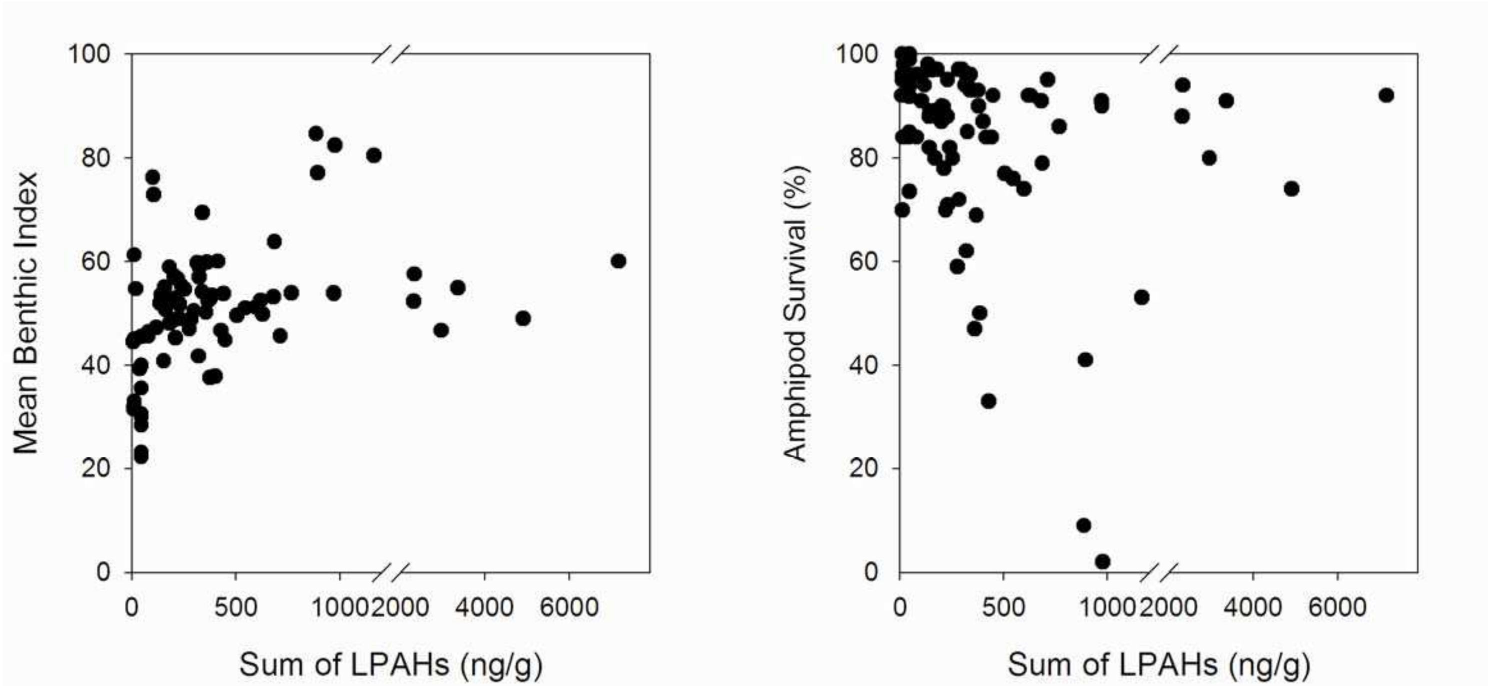

Figure 17. Relationships between LPAHs : mean benthic index, and LPAHs : control-adjusted percent amphipod survival in sediments from clean-up areas (n = 87).

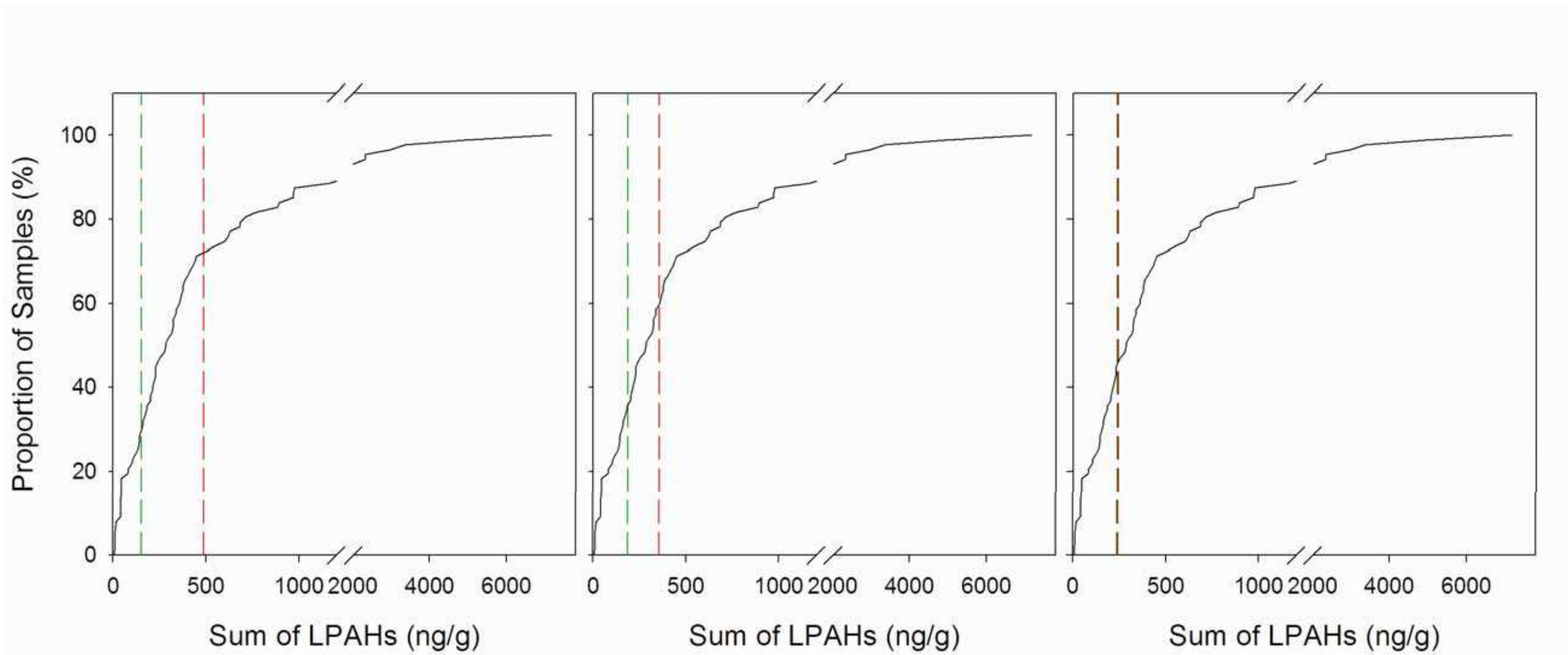

Figure 18. Cumulative distribution function (CDF) for LPAHs in sediments from clean-up areas (n = 87). Dashed lines indicate prediction limits for 5 future samples at probability levels of 95% (left plot), 90% (center plot), and 80% (right plot). Red lines are upper limits calculated from un-impacted samples (SQO scores = 1 – 2). Green lines are lower limits calculated from impacted samples (SQO scores = 3 – 5). The proportion of future samples below the red line are representative of the population of un-impacted samples with 95%, 90%, or 80% statistical confidence, whereas the proportion of future samples above the green line are representative of the population of impacted samples at that statistical level.

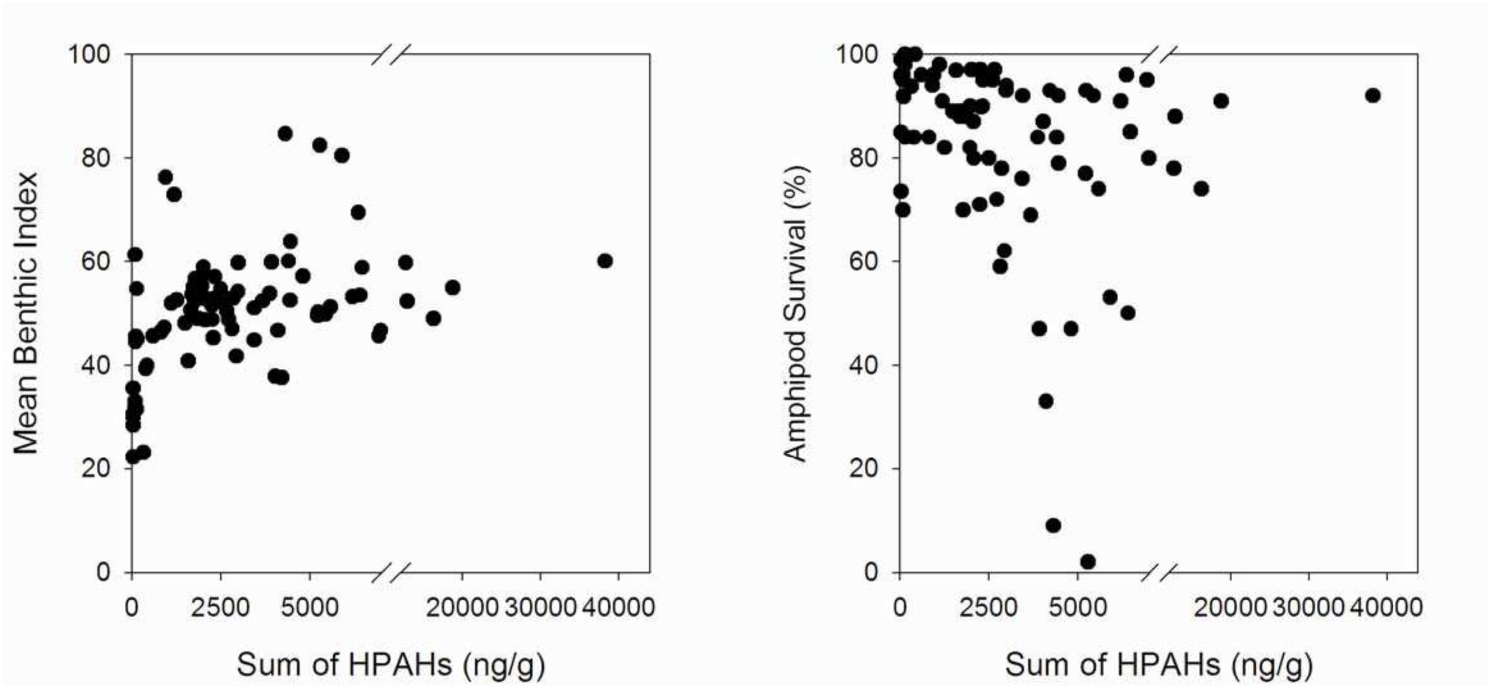

Figure 19. Relationships between HPAHs : mean benthic index, and HPAHs : control-adjusted percent amphipod survival in sediments from clean-up areas (n = 87).

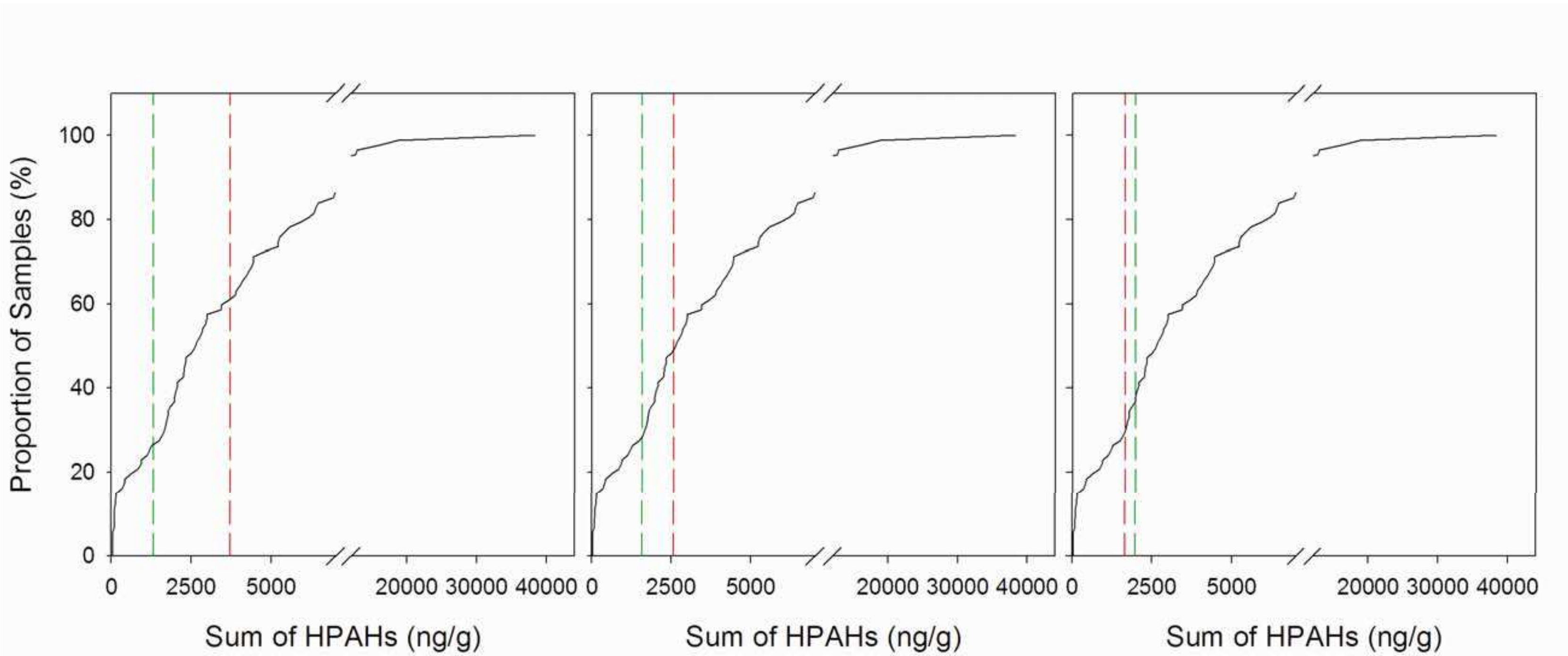

Figure 20. Cumulative distribution function (CDF) for HPAHs in sediments from clean-up areas (n = 87). Dashed lines indicate prediction limits for 5 future samples at probability levels of 95% (left plot), 90% (center plot), and 80% (right plot). Red lines are upper limits calculated from un-impacted samples (SQO scores = 1 – 2). Green lines are lower limits calculated from impacted samples (SQO scores = 3 – 5). The proportion of future samples below the red line are representative of the population of un-impacted samples with 95%, 90%, or 80% statistical confidence, whereas the proportion of future samples above the green line are representative of the population of impacted samples at that statistical level.

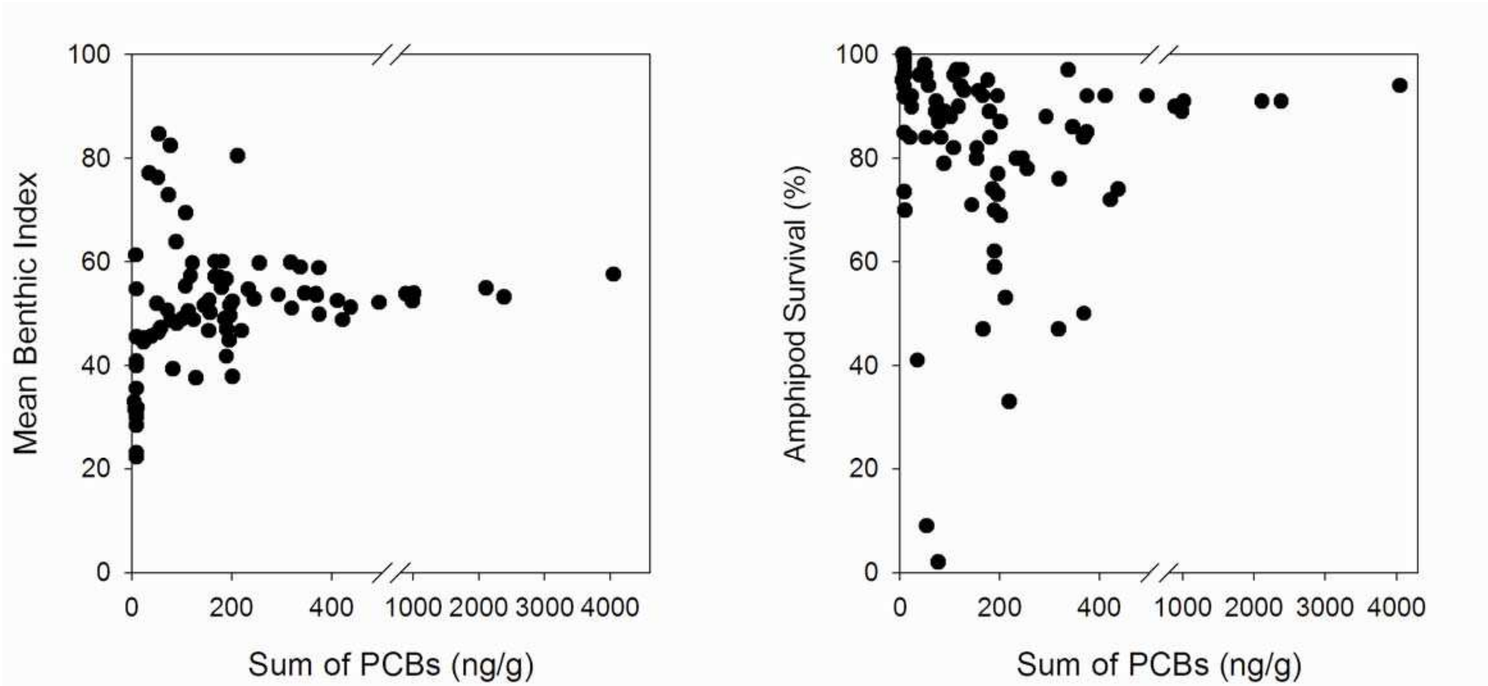

Figure 21. Relationships between PCBs : mean benthic index, and PCBs : control-adjusted percent amphipod survival in sediments from clean-up areas (n = 87).

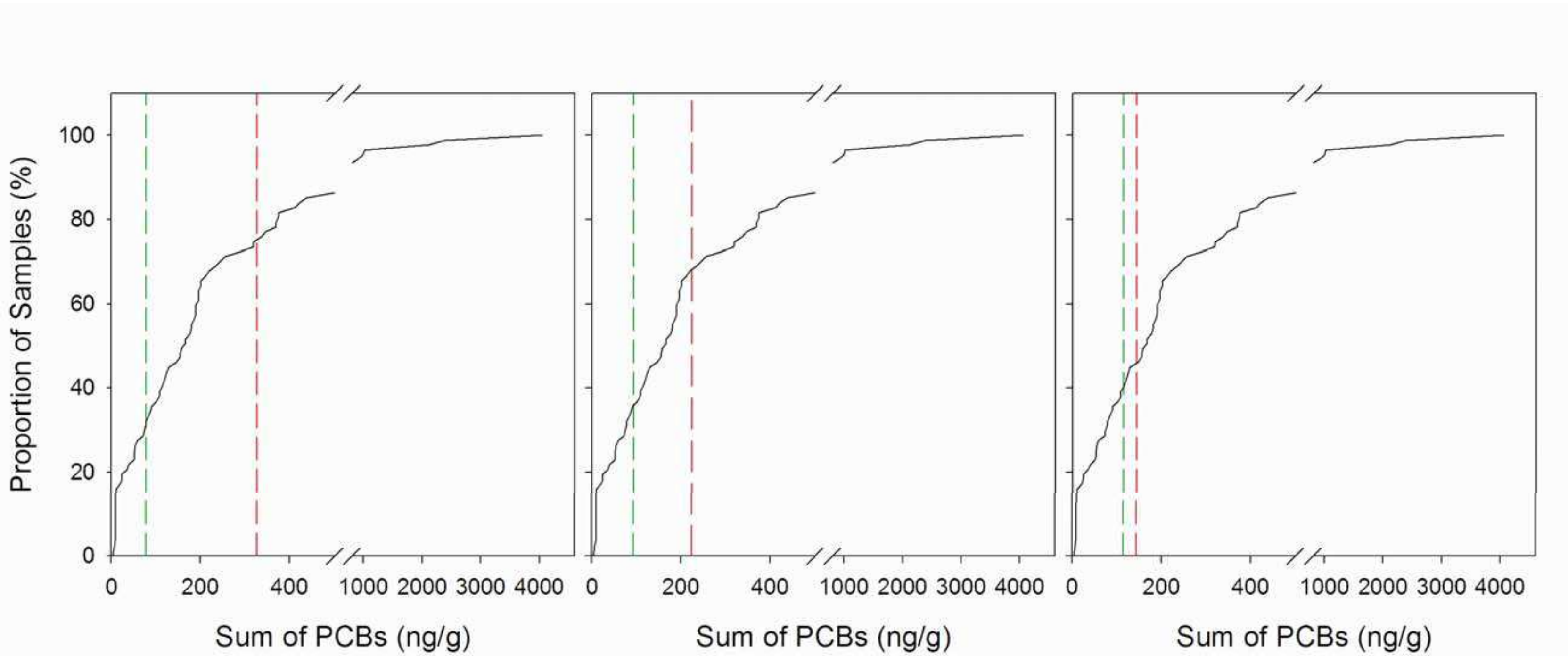

Figure 22. Cumulative distribution function (CDF) for PCBs in sediments from clean-up areas (n = 87). Dashed lines indicate prediction limits for 5 future samples at probability levels of 95% (left plot), 90% (center plot), and 80% (right plot). Red lines are upper limits calculated from un-impacted samples (SQO scores = 1 – 2). Green lines are lower limits calculated from impacted samples (SQO scores = 3 – 5). The proportion of future samples below the red line are representative of the population of un-impacted samples with 95%, 90%, or 80% statistical confidence, whereas the proportion of future samples above the green line are representative of the population of impacted samples at that statistical level.

APPENDIX I. Mean and standard deviations for each biological indicator at two spatial scales: 1) all San Diego Bay, and 2) all clean-up areas.

Chemical/Parameter	n	Mean	S.D.
All San Diego Bay			
Amphipod Survival (%)	161	84.5	16.2
Mean Benthic Index	161	44.4	13.7
Benthic LOE	161	2.2	0.8
Toxicity LOE	161	1.9	1.0
SQO Score	161	2.7	1.4
Clean-up Areas			
Amphipod Survival (%)	87	82.5	18.4
Mean Benthic Index	87	51.2	11.7
Benthic LOE	87	2.5	0.7
Toxicity LOE	87	1.9	1.0
SQO Score	87	3.1	1.4