N00102.AR.000597
NSY PORTSMOUTH
5090.3a

DEPARTMENT OF THE NAVY
NORTHERN DIVISION
NAVAL FACILITIES ENGINEERING COMMAND
10 INDUSTRIAL HIGHWAY
MAIL STOP, #82
LESTER, PA 19113-2090

IN REPLY REFER TO
5090
Code 1823/FE

25 SEP 1998

Ms. Meghan Cassidy
U.S. Environmental Protection Agency
Region 1
JFK Federal Building HBT
Boston, MA 02203-2211

Mr. Iver McLeod
Maine Department of Environmental Protection
State House Station 17
Augusta, ME 04333-0017

Subj: OFFSHORE ECOLOGICAL RISK ASSESSMENT; PORTSMOUTH NAVAL SHIPYARD, KITTERY, ME

Dear Ms. Cassidy and Mr. McLeod:

Enclosed please find our responses to comments on the Maine Department Of Environmental Protection's comments dated March 9, 1998, and the Lepage Environmental Services comments dated March 23, 1998, and David R. Brown, SC.D. comments dated March 13, 1998. We are also including a copy of the revised Appendix IX paper for the Benthic Assessment Endpoint, a copy of the "reader friendly" Executive Summary, and the Glossary. Your comments are r quested on or before October 26, 1998.

If additional information is required, please contact me at (610) 595-0567, x159.

For the Community Restoration Advisory Board (RAB) members; if you have any comments or questions on these issues, they can be provided to the Navy at a RAB meeting, by calling the Public Affairs Office at (207) 438-1140 or by writing to:

Portsmouth Naval Shipyard
Code 106.3 Building 44
Attn Ms Marty Raymond
Portsmouth, NH 03804-5000

Sincerely,

Frederick J. Evans

FREDERICK J. EVANS, P.E.
Remedial Project Manager
By direction of the
Commanding Officer

Encl:
(1) Response to Comments
(2) Revised Appendix IX paper
(3) Executive Summary
(3) Glossary

Subj: OFFSHORE ECOLOGICAL RISK ASSESSMENT; PORTSMOUTH NAVAL SHIPYARD, KITTERY, ME

Distribution:

NOAA (K. Finkelstein)
ME Dept. of Marine Resources (D. Card)
PNS (Code 106.3R, M. Raymond)
MESO (R. Johnston)
Ms. Juanita Bell
Mr. Jeff Clifford
Ms. Eileen Foley
Mr. Phil McCarthy
Mr. Guy Petty
Mr. Peter Vandermark
Brown and Root Environmental (L. Klink, B. Horne)
US Fish & Wildlife Service (K. Munney)
NH Fish & Game (J. Nelson)
COMSUBGRU TWO (R. Jones)
UNH JEL (F. Short)
Mr. Doug Bogen
Ms. Michele Dionne
Ms. Mary Marshall
Mr. Jack McKenna
Mr. Onil Roy
PNS (100PAO) w/o encl

7. Risk Characterization

Based on the approach laid out by the U.S.EPA Risk Assessment Forum (U.S. EPA 1992, 1996a), two types of information are required to characterize ecological risks. These are data on the exposure in environmental media (river water and sediments) and data that relate exposure levels (dose) to measurable effects. The analysis activities that were conducted to characterize risk consisted of evaluating the weight of evidence of risk for each AOC for the assessment endpoints of concern, analyzing bioaccumulation data, and identifying potential risk drivers. Developed by the Massachusetts Weight-of-Evidence Workgroup (Menzie *et al.* 1996), a weight-of-evidence approach was used to evaluate measures of effects and measures of exposure to interpret the level of risk evident for each applicable assessment endpoint and AOC. To develop risk characterization narratives for each AOC, the weight-of-evidence information was summarized. The bioaccumulation data were analyzed to relate exposure conditions to tissue concentrations of receptors of concern and, to the extent possible, assess ecological effects. By relating evidence of risk to exposure pathways, the ecological risks associated with environmental media in each AOC were identified. Using a probabilistic approach to identify COCs which could be potential risk drivers, media-specific exposure-point concentrations were used to estimate the probability of exceeding an benchmark, criteria, or standard. For effects endpoints with continuous dose-response relationships the probability of an effect was simulated. In addition, potential linkages to the SWMUs were identified and the uncertainties associated with risk characterization activities were identified and discussed.

7.1 Weighing the Evidence of Risk

7.1.1 Rationale

A weight-of-evidence procedure was used to relate multiple measures of exposure and effects to the assessment endpoints used in the ecorisk assessment for Portsmouth Naval Shipyard (Munns *et al.* in prep.). Risk characterization consists of drawing conclusions about the ecological risks associated with environmental stressors from data obtained from field and laboratory studies. Results from multiple measures of environmental conditions must be synthesized and reconciled. Because there may be conflicting or ambiguous results, a large degree of professional judgment is required to interpret risk and define the degree of environmental harm. Recently, the Massachusetts Weight-of-Evidence Workgroup (Workgroup) developed an approach to formalize professional judgment and reach conclusions about ecological harm by following a weight-of-evidence procedure (Menzie *et al.* 1996). The Workgroup's approach involved: assigning weights to each measurement endpoint, summarizing the responses observed in each measurement endpoint, and evaluating the degree of concordance in the responses of measurement endpoints.

Using professional judgment to establish measurement weights and to interpret endpoint response, a team of risk assessors and risk managers (Ecorisk Team) applied the Workgroup's approach to characterize ecological risk from past hazardous waste practices at the Portsmouth Naval Shipyard. This highly interactive, consensus-building process led to refinements and extensions of the methods offered by Menzie et al. (1996). By taking into account the strengths and weaknesses of the measurement methods used, the weight-of-evidence approach was applied

For the weight-of-evidence analysis, the outcome (interpretation) from the individual measures was used, not the actual results. This can be made clear by reviewing the discrete steps used in the procedure:

1. Endpoint weights were objectively assigned to each measure of exposure and effect. The endpoint weight was based on the strength of the relationship to the assessment endpoint, data quality, and study design.

2. The outcomes of the measures were interpreted based on whether the result added weight to the conclusion of risk or no risk (see Table 7-1 for scheme used to interpret the outcomes of measurement activities). Summary tables for each assessment endpoint and AOC were constructed that contained all the information available to evaluate risk (these are the tables provided in Appendix IX).

3. Definitions of risk were developed to interpret the results of the exposure and effects information (see pp 7-5 to 7-6 and Table 7-2).

4. Scatter plots of the outcomes of the exposure and effects measures were plotted versus their corresponding endpoint weights. This allowed the results obtained for each assessment endpoint and AOC to be visualized.

5. A centroid was calculated that consisted of a weighted-average of the outcomes (weighted by their endpoint weights). The centroid was used to aid in interpreting the balance of exposure and effects information suggested by the data. Higher weighted measures would tend to draw the centroid in their direction.

6. Conclusion about risk were developed based on the outcomes and confidence levels shown in the scatter plots.

While the results of the individual measures may not be linear, the outcomes of the measurements are linear. For example (from Table 7-1), negligible exposure < low exposure < elevated exposure < high exposure < adverse exposure, and no effect < potential effect < probable effect. Therefore, weighing the evidence is not canceling out risks, rather it is developing a conclusion that reconciles the information available.

7.1.3 Weight-of-Evidence Procedure

A methodology for evaluating multiple lines of evidence from endpoints measuring environmental harm was recently proposed by the Massachusetts Weight-of-Evidence Workgroup (Menzie *et al.* 1996). This weight-of-evidence approach involves steps of weighing endpoints, evaluating endpoint responses, and evaluating of the degree of concurrence among multiple endpoints. This methodology was applied in the assessment of ecological risks to estuarine systems associated with waste management practices at the Portsmouth Naval Shipyard in Kittery, ME.

Following the approach recommended by the Workgroup (Menzie *et al.* 1996) exposure and effects measures were weighted based upon several attributes falling into three general categories:

concern. The weight-of-evidence analysis evaluated all data for each AOC relative to the potential to cause ecological effects. The most conservative level of exposure present within the AOC was used to develop the conclusions of negligible, low, elevated, high, or adverse exposures, therefore the weight-of-evidence procedure did not underestimate exposure. While there were low sample sizes for some media in a few of the areas of concern, the weight of evidence was based on all available data and no samples were eliminated due to low sample size.

The interpretation of effect measurement outcomes ranged from no effect to probable effect (Table 7-1). Differences among the descriptors used to interpret exposure and effect outcomes reflected the opinion that concentrations of COCs could be measured directly and reasonably accurately, whereas linkages between observed effect and exposure to COCs could only be inferred.

The first two steps of the weight-of-evidence process culminated in a set of tables summarizing the weights, outcomes, and interpretations of measures of exposure and effect (Appendix IX). Recognizing that confidence in determinations of risk would be bolstered by the cooccurrence of the stressor and ecological response (U.S. EPA, 1992), measures were grouped into exposure and effect categories. Organized by area of concern and assessment endpoint, the summary tables also identified issues that were felt to be significant with respect to understanding the uncertainties associated with the evidence and its interpretation. Together with supporting narrative (see below), the summary tables enhanced the transparency of the weight-of-evidence process to risk managers and other stakeholders by laying out the logic behind evaluation of exposure and effect information. Summary tables developed for the each area of concern and assessment endpoint are provided in Appendix IX.

In order to interpret outcomes of exposure and effects measures in terms of risk, risk definitions were developed. The risk definitions considered the following factors:

- The amount of evidence (preponderance)
- The degree of exposure or effect evidence (magnitude)
- The spatial extent of measured impacts (extent)
- The strength of relationship between exposure and effects information (causation)

The definitions of risk were used to interpret the results of exposure and effect evidence (Table 7-2, see Section 8.2 for discussion of the risk management options associated with the interpretations of risk). Resulting from (1) no evidence of effect and negligible or low evidence of exposure or (2) some evidence of potential effects and negligible evidence of exposure, negligible risk suggested no impacts based on exposure and effects measures. Negligible risk was associated with the general lack of evidence of exposure and effects in the field data.

A determination of low risk resulted from (1) no evidence of effect and elevated or high exposure, (2) evidence of potential effect and low exposure, or (3) evidence of probable effect and negligible or low exposure. Based on the field measurements, low risk typically lacked demonstrable correlations between exposure and response. Low risk suggested limited impacts based on exposure and effects measures.

For intermediate risk there were (1) no evidence of effects and evidence of adverse exposure levels, (2) evidence of potential effects and evidence of elevated to high exposure, or (3)

objectives, the data were not used in the risk characterization. During the *ECOS* cruise conduced July 1992, continuous measurements of dissolved oil[2] and ultra-trace analysis of discrete water samples for organic compounds[3] did not detect any instances when water quality criteria were exceeded. During the three *Gulf Challenger* cruises conducted fall 1993, dissolved and particulate metal concentrations were measured throughout the lower estuary (and during third cruise the whole estuary). Data on COC concentrations were only obtained for Cd, Cu, Ni, and Pb (data for Fe, Mn, Al, salinity, and total suspended solids were obtained as well as). In the weight-of-evidence procedure these data were used as indicators of water quality for all the COCs.

For phytoplankton biomass, data quality was high, but the strength of relationship and study design were low. This is because the high amount of natural variability and low sampling frequency resulted in the data being unable to discriminate between contaminant responses and natural variability. For the water benchmark measures, high data quality and medium strength of relationship and study design were assigned to the measures of esturaine surface water while the seep samples were rated low for strength of relationship and study design. This was because the seeps would be rapidly diluted upon entering in the harbor. Unless sampled frequently, water chemistry data are subjected to rapid spatial and temporal changes. The strength of relationship for sea urchin toxicity was high, but the holding times were exceeded and the test was only conducted on one occasion. An overall medium weight was assigned to deployed mussel condition and tissue residues. The condition measure (Scope for Growth) is highly correlated with somatic growth, however it may not be as sensitive as growth measurements. There is good linkage between the deployed mussel tissue residue levels and effects from bioaccumulation in the assessment endpoint (Appendix IX.1).

Measurements of winter flounder abundance, size, histopathology and tissue residues were evaluated for Portsmouth Harbor as a whole. While demersal fish such as winter flounder are good indicators of environmental pollution because of their relatively long life span, their association with bottom sediments, and the fact that they feed mainly on benthic invertebrates, they are migratory and they may not stay in close proximity to site. Winter flounder can inhabit virtually any area of the estuary but quantitative sampling in the Piscataqua River is difficult because of the many mooring, lobster pots, rockpiles, and other obstructions. The measures of abundance and size were assigned an endpoint weight of low and the measures of histopathology and tissue residues were assigned an endpoint weight of medium (Appendix IX.1).

7.1.4.3 Epibenthic Assessment Endpoint

To assess effects on epibenthic receptors, lobster density, indigenous mussel density, shell length, and condition index, and fucoid biomass were evaluated. To assess exposure to epibenthic receptors, estuarine surface water concentrations, seep water concentrations, and tissue residues of lobsters, fucoids, and mussels were evaluated (Appendix IX.2).

---

[2] Dissolved oil was measured by continuously monitoring oils fluorescence with a fluorometer calibrated with dissolved diesel fuel.

[3] PAHs, PCBs and pesticides were below detectable levels in these samples.

judging environmental harm considering that major physical characteristics such as current velocity and substrate type have a greater influence on algal presence and abundance. Fuciod tissue residues were rated higher because there may be a better relationship between contaminant exposure and uptake, particularly, with heavy metals and this has been documented in the scientific literature. Sampling was only conducted once and limited to areas where fucoid habitat was present (5 of the 7 areas of concern). There was substantial variability associated with the data. "The pronounced variability of Ascophyllum's biomass is caused, at least in part, by spatial variability of tidal current regimes within this section of the Piscataqua River" (Mathieson 1994). No statistical analysis was performed to determine significant results between the reference location and areas of concern..

Measures of mussel density, condition index, shell length, and mussel tissue residues were all assigned a medium weight. The mussel, *Mytilus edulis*, is a sessile filter-feeding bivalve known to accumulate contaminants in the water column and is a good indicator for evaluating water quality. Many studies have supported the relationship between surface water contaminant concentrations and mussel tissue residue but mussel density measures alone have limited utility in determining whether surface water contaminants attributable to the Shipyard adversely impact the mussel population. Mussel condition index and shell length may relate to the success of the population and the absence of larger mussels may indicate that an area is exposed to a stressor. However differences in size distributions could also due to other factors such as current velocity, predation, and ice scouring. The sampling of mussel effects measures only took place once (fall 1991) at 24 stations including reference areas, but no confirmational sampling was conducted. The results were quantitative, they were tested for statistical significance, and a standard method was employed. Abundance measures were highly variable and are very likely subjected to physical factors rather than from exposure to contaminants (Appendix IX.2).

The measure of exposure from mussel tissue residues is well established and has been used in NOAA's Status and Trends Program and in EPA's Mussel Watch Program. The State of Maine has also been collecting mussel tissue data since 1987. Statistical analyses were conducted to identify statistically significant contaminant accumulation in mussel tissues. A significant correlation ($p<0.01$) between mussel tissue residues and sediment contaminant concentrations for lead and copper was reported. Some of the stations where elevated mussel tissue contaminant concentrations were detected also had samples with elevated metals in eelgrass leaves. The consensus from the WOE analysis was that mussel tissue residues should be assigned a "Medium" strength of relationship because ecological effects from tissue concentrations can not be inferred. For the mussel residue data there was good spatial coverage and mussels were sampled on a quarterly basis for two years. However, the time period was too short to determine seasonal trends and it is difficult to compare analytical results between Phase I and II (transect and point sampling) since sampling stations and methodologies were different (Appendix IX.2).

Measures of exposure obtained from estuarine surface water and seep benchmarks were similar to Pelagic Assessment Endpoint discussion, except for the seeps. For the epibenthic assessment endpoint there is a stronger association between the seeps and its potential effect on the blue mussel and fuciod algal communities. This is because seeps have been observed to discharge in areas where blue mussels and algae are present. Therefore, measurement of the seeps was ranked as medium (Appendix IX).

organic carbon ~ 2%. The U.S. EPA Atlantic Ecology Division has an extensive database of no toxicity associated with this reference sediment. A concurrent water only reference test was also conducted during the test to fulfill quality assurance/quality control requirements. There is a direct biological relationship between the results of the sediment toxicity test and the assessment endpoint since the test species, *Ampelisca abdita* is one of the dominant taxa identified during the benthic community analysis. The test was only performed once (Fall 1991 - Phase I) but benthic community analysis, analysis of sediment characteristics and chemical analysis were performed concurrently on the 23 Phase I sampling locations.

Comparison of sediment chemistry data to crustal ratio received a medium weight. There is no biological linkage between crustal ratio and infaunal community. "The degree of enrichment, or deviation from the expected metal concentration, indicates that there could be alternative sources of metal contributing to the observed distribution, presumably anthropogenic, but also possibly involving local geological mineral inputs or the atmospheric fall out of dust particles" (Johnston et al. 1994c). There is little correlation between chemical stressors originating from the shipyard and infaunal community response based on crustal ratios. Crustal ratio was integrated into a ranking scheme utilized to identify the most important contaminants of concern and it's application is limited to metals. Metal enrichment could be a result of historic releases from the shipyard.

Measures of exposure determined from bulk-sediment chemistry concentrations received a medium weight. Bulk-sediment concentrations were compared to NOAA ER-L and ER-M values and Washington State Cleanup Levels (WA-CL). The consensus from the WOE analysis was that bulk-sediment chemistry should be evaluated as one measure of exposure and the magnitude of contamination should was evaluated as follows:

≤ reference condition or below conservative benchmark (ER-L) = negligible exposure

> qualitative screening level = low exposure

statistically > reference condition = elevated exposure

> a conservative benchmark (ER-L) = high exposure

> a nonconservative benchmark (ER-M or WA-CL) = adverse exposure.

Based on the NOAA database, the NOAA ER-L concentration for a specific contaminant represents the overlap of various biological effects data (lab tests, field tests and modeling) with chemical concentrations. The ER-L represents the 10th percentile of data that demonstrated a biological response. The range of contaminant concentrations between the ER-L and ER-M represent possible effects, "The range in concentrations over which toxic effects are occasionally observed" (Long et al. 1995). The response is not well correlated with duration of exposure and this measure does not take into account bioavailability or bioaccumulation. Contaminant concentrations in sediments that exceed the ER-M may be frequently responsible for toxic effects. "The concentrations equivalent to and above the ER-M value represent a probable-effects range within which effects would frequently occur" (Long et al., 1995). There is no biological relationship between the WA-CL approach and the infaunal community at Portsmouth Naval

concentrations for eelgrass receptors, the fact that a correlation between stressor levels and eelgrass effects measures have not been demonstrated, and that the scientific basis for inferring environmental harm from the eelgrass measurements is weak. Additionally, no data are presently available to relate exposure measures (water, sediment, and tissue concentrations) to biological effects to eelgrass.

The measure of effect based on the spatial distribution of eelgrass was assigned a medium endpoint weight. The spatial distribution of eelgrass beds has a strong association with the assessment endpoint because the presence or absence of eelgrass is related to eelgrass health and vitality. The correlation between stressor levels and the spatial distribution of eelgrass is expected but has not been demonstrated. Benchmarks for chemical effects on eelgrass spatial distribution are not available, therefore environmental harm is assumed. The sampling design used to measure eelgrass distribution had low weight, particularly in discriminating responses from contaminants. However, qualitative information on the location and distribution of eelgrass beds in the lower estuary was obtained. The data were collected from only one sampling period (fall 1991) and the data were derived subjectively. Although the method was not previously published, its suitability and applicability are documented (Mueller et al. 1992, Short 1994a).

The measures of effect based on eelgrass leaf and root/rhizome morphology were assigned an endpoint weight of medium. The measures of morphology have a direct link to the eelgrass assessment endpoint, however, correlation between the measurement and stressor levels has been suggested but not proven. Eelgrass root+rhizome biomass was negatively correlated with sediment Pb concentration and a cause-and-effect has been suggested but not proven and the scientific basis for inferring environmental harm is weak. The morphology study was designed to determine patterns of plant morphology and the study provided direct information on the status of eelgrass morphology in the estuary. The data were only collected from one sampling period (fall 1991) and study was descriptive, so no statistical or biological significance could be interpreted (Appendix IX.3). The method has been published (Mueller et al. 1992).

Concentrations of contaminants in eelgrass leaves and root/rhizome were assigned a medium endpoint weight. Metal concentrations in the plants are biologically mediated however, no acceptable benchmarks for inferring environmental harm are available. The eelgrass tissue cut-off (ETC) was used in this study. The ETC, defined as the mean of the reference site concentrations plus 3 standard deviations, was used as an indicator to identify samples significantly above background level. The ETC is a personal index and has limited applicability or certainty and the scientific basis for inferring environmental harm is weak and relatively insensitive. There was a quantitative correlation between the eelgrass tissue concentrations and stressor levels, however, biological significance has not been demonstrated. Tissue concentrations in eelgrass root/rhizome samples were positively correlated with sediment chemical concentrations, but statistical significance was not demonstrated. There is an expected relationship between eelgrass tissue concentrations and biological effects to eelgrass, however, no data are available to support this assumption. Metal concentrations in eelgrass leaf and root tissues were analyzed for the initial baseline data collected in September 1991 and then quarterly starting in December of 1991 through September of 1992. Station 1 (Portsmouth Harbor), and Station 23 (York River) were the two reference stations sampled for eelgrass. Also multiple exposure measures were sampled during appropriate periods, however, not enough sampling was

(York River Estuary). Reference marshes included fairly well developed marshes on the southeastern shore of Shapleigh Island in Little Harbor, another south of Admiralty Village on Spruce Creek. The methodology produced site specific information related to the assessment endpoint, but the methods were subject to a high degree of natural stochasticity. The sampling consisted of only one sampling period (Oct. 23- Nov. 5, 1922), quantitative data were obtained, and the data were statistically tested, but biological significance can not be inferred (Appendix IX.5) .

The measure of exposure from metal concentration measured in *Spartina* leaf tissue was assigned a medium endpoint weight. The issue of metal bioavailability and bioaccumulation was not within the scope of the study, so only 14 composite samples of *Spartina* leaf tissue were analyzed. Currently, no benchmark or warning levels are available for tissues of salt marsh plant, so exploratory analysis focused on patterns of metals within and between marsh locations. There was a positive correlation between salt marsh tissue concentrations and chemical stressor levels measured in sediment samples, but statistical significance has not been demonstrated. Although, a relationship between exposure measures and biological effects to *Spartina* are expected, no data are available to support this assumption. Measures of *Spartina tissue* concentrations are directly related to stressor levels and there were medium variability and uncertainty in the data. Quantitative exposure measures were tested statistically, but biological significance cannot be clearly inferred. The analytical method has not been accepted by regulatory authorities, but it has been published in peer-reviewed literature (Appendix IX.5).

The measure of exposure from bulk-sediment chemical concentrations was assigned a low endpoint weight. This was because benchmarks which can be used to relate sediment concentrations to effects to salt marsh receptors are not available. The degree of exposure from bulk-sediment chemical concentrations was evaluated in the same manner as was used for the benthic assessment endpoint (see Section 7.1.4.4).

7.1.4.7 Avian Assessment Endpoint

Potential exposure to avian receptors was evaluated by evaluating the modeled dose from dietary exposure to toxicity reference values (TRVs) developed for black ducks (omnivore), Canada geese (herbivore), herring gulls (carnivore), and ospreys (piscivore). No direct measures of effects of avian receptors were made during the estuarine study (Appendix IX.6). Because avian receptors would utilize all habitat areas in the lower estuary, exposure levels were evaluated for Portsmouth Harbor as a whole (Portsmouth Harbor focus area). This was accomplished by evaluating the maximum exposure levels from food (prey and plants) and sediment in the model calculations.

The data used to evaluate exposure to avian receptors had high data quality. Major sources of uncertainty in the assessment of exposure to avian receptors were association with the assumptions used in the model (see section 6.0, Appendix IX.6). Since direct measures of effects were not made, effects were assumed if the dietary hazard quotient (HQ) HQ >> 1 or hazard index (HI) HI > 10 (Ken Munney, USFWS, personal communication).

Avian species were determined to be useful indicators of potential impacts to upper food chain biota for PNSY. However, assessment and measurement endpoints were not originally selected to include representative indicator avifauna. Therefore, assessment of potential upper

Summary tables (7-4 to 7-9) of the weight of evidence were constructed with the entries obtained from scatter plot of effects and exposure measures versus their endpoint weights. The level of risk indicated from the interpretation of exposure and effect evidence in determining risk (see Table 7-2) and the confidence associated with the determination of risk were also included in the tables. The confidence level was based on the following factors:

- the average of the endpoint weights for effects and exposure measures,
- the degree of concurrence among the weights (e.g. scatter of weights), and
- the degree of concurrence between conclusions regarding magnitudes of exposure and effect (e.g. the balance between the average endpoint weight and the scatter of weights column).

Where necessary, professional judgment was used to qualify the conclusions.

7.1.5 Clark Cove

Located near the center of Portsmouth Harbor and adjacent to the Jamaica Island landfill, Clark Cove was a major focus of study during the ecorisk assessment. Formed by filling the area between Seavey Island and Jamaica Cove and the construction of the causeway to Clark Island, Clark Cove is a major depositional area in the lower estuary. Due to its physical configuration water movement within the cove is very restricted. This results in the accumulation of mud and muddy-sand deposits in the cove. The bottom topography of the cove, altered by dredging to form the sounding basin, makes the cove even more conducive to the accumulation of sediment. The shoreline of the cove has been highly modified by the causeway and the shoring up of the beach face along the landfill with riprap, cap, and fill material. The protected harbor afforded by the cove is currently used as a marina and many boats are moored in the cove from late spring to early fall. Ecologically the cove provides habitat for a wide variety of fish and invertebrates (including winter flounder and lobster), extensive mussel beds were located within the cove, some mudflats were present along the southern edge of Jamaica Island, and rockweed algae were present along the exposed rocky intertidal shoreline. In the area along the northeastern edge of the cove where bottom sediment has remained undisturbed, an established eelgrass bed is present. A small marsh with well-developed peat substrata is also present along the northwestern side of Clark Island.

7.1.5.1 Clark Cove Pelagic

Indicating a probable effect, statistically significant toxicity to sea urchin fertilization was observed at three of the six stations sampled in Clark Cove. Data from the SFG measurements conducted on Phase I deployed mussels and monthly phytoplankton biomass measurements indicated no effect because the data from Clark Cove were similar to reference areas. Since phytoplankton biomass had a low endpoint weight and SFG and sea urchin toxicity both had medium endpoint weights, the weighing of effects measures resulted in a conclusion of potential effects with a medium level of confidence (Figure 7-1). High exposure was suggested from seep samples and deployed mussel samples. The seep samples exceeded water quality criteria for Cu, Ni, and Zn. The deployed mussel samples from Phase II had Pb above predeployment levels but

effects to epibenthic receptors. Confidence was increased by the relatively large sample sizes and moderate sampling and analytical variability in the epibenthic measurement data for Clark Cove.

7.1.5.3 Clark Cove Benthic

A potential effect to species evenness was suggested at two (stations 4 and 8) of the six stations in Clark Cove (which had greater than the 95th percentile of the evenness index– density/taxa). The other benthic measures – invertebrate density, species richness, and amphipod toxicity – showed no effect and the conclusion was a high confidence of no effect (Figure 7-3). Two of the stations in Clark Cove (stations 5 and 7) had relatively low infauna density and the number of taxa were the lowest measured. However, pollution-sensitive ampeliscids were present at both of the stations and the density and richness values were above the 5th percentile determined for the Piscataqua River distribution. There was a medium confidence of elevated exposure to benthic receptors. This was supported by adverse exposure, indicated by bulk-sediment concentrations of Cu above the ER-M level, elevated exposure levels suggested by enrichment of Pb, Hg, Cr, Cd, As, and Ag in the sediment, and negligible exposure indicated by SEM-AVS and predicted EqP pore water concentrations (Figure 7-3).

Issues associated with effects and exposure measures used to assess the benthic community for Clark Cove were identified. There was no data on the seasonal variation of sediment AVS-SEM concentrations. The effect that the fine texture of sediment substrate had on obtaining accurate samples from the cove and the degree to which the sampling locations were representative of benthic conditions in the cove were also considered. There was only a moderate level of sampling and analytical variability in the analytical chemistry data but there was a high degree of variability in the benthic community data from Clark Cove. Three of the stations had very high densities of benthic invertebrates while two of the stations had very low densities and species richness ranged from relatively high to very low. There was high enrichment of metals in the sediments of the cove (Pb, Hg, Cr, Cd, As, and Ag), many chemicals (PAHs, pesticides, tPCB, and metals) exceeded ER-L thresholds, and the ER-M threshold was exceeded by Hg, tPCB, and pesticides. However, no COCs exceeded pore water toxicity levels, SEM metals were far below AVS levels, and the sediments of the cove were not toxic to amphipods.

7.1.5.4 Clark Cove Eelgrass

For the eelgrass bed located on the northeastern edge of the cove, no indication of effect was detected by any of the morphological or density measures. Leaf length and biomass, root/rhizome length and biomass, vegetative shoot density, and reproductive shoot density were above or within the range measured for eelgrasses sampled from reference locations. Because eelgrass beds were absent from inner Clark Cove, the absence of eelgrass was interpreted as a potential effect. Since most of Clark Cove is too deep to support eelgrass growth, the extent of the affected area was limited to suitable eelgrass habitat along the fringes of the inner cove (Hoven and Short 1996). For exposure measures, both leaf and root tissue were elevated above pristine conditions identified by the eelgrass tissue cutoff (ETC) level for some metals and bulk-sediment concentrations sampled from the eelgrass bed (station 3) exceeded the ER-L thresholds for As, Cr, Hg, and Ni. The findings were characterized as medium confidence of potential effect and medium confidence of elevated exposure to eelgrass receptors (Figure 7-4).

7.1.6.1 Sullivan Point Pelagic

Effects to pelagic receptors in the Sullivan Point AOC were assessed by measurements of toxicity to *Arbacia* fertilization which showed no toxicity in the water samples collected from a single station offshore of the point (station 9, n=4 replicates). For exposure measures, high exposure was suggested by seep water samples that exceeded ambient WQC for Cu, Hg, and Zn, while negligible exposure was indicated by estuarine surface water concentrations that were well below ambient WQC. The conclusion reached was that there was medium confidence of no effect and low confidence of low exposure to pelagic receptors at Sullivan Point (Figure 7-6).

The issues associated with the pelagic measures for Sullivan Point were that the toxicity test was the only measure of effect and the holding times were exceeded for the samples. No mussel deployments were conducted for this AOC. There was good agreement between the lack of exposure levels measured in estuarine surface water and the lack of toxicity observed in the water samples. Another source of uncertainty was that the seep discharge into Sullivan Point was not well characterized (six samples from two stations from the April 1993 sampling period) and the rate of seep flow and how long it would take to for the seeps to be diluted is not known.

7.1.6.2 Sullivan Point Epibenthic

All the measures of effects to epibenthic receptors in the Sullivan Point AOC showed no effect (Figure 7-7). Fucoid biomass, mussel condition index, mussel density, and mussel shell length were all within the range measured at reference locations. High exposure was indicated from the seep water data. Elevated exposure was suggested by the juvenile lobster tissue data that were significantly higher than reference concentrations for PAHs and Hg and by mussel tissue concentrations that were above warning levels for Pb and Cr. Low exposure was indicated from fucoid alga tissue concentrations that were above average for Cu. Negligible exposure was indicated by estuarine surface water concentrations that were far below ambient WQC levels. These data resulted in the conclusion that there was medium confidence of no effect and elevated exposure to epibenthic receptors in the Sullivan Point AOC (Figure 7-7).

Issues associated with epibenthic measures for Sullivan Point were identified. There was substantial variability in fucoid biomass and high Cu concentrations in fucoid tissues collected at this location. There was not very much information about the seeps. The route of exposure and uptake of the elevated residues measured in the lobsters and mussels could not be identified. There was also uncertainty about whether the elevated tissue residues could affect the lobsters or mussels. No effects to fucoids or mussels were detected at this location.

7.1.6.3 Sullivan Point Benthic

Measures of effects to benthic receptors in the Sullivan Point AOC showed no effect from the similarity of the benthic community's density, species richness, and species evenness to reference areas. However, because a probable effect was indicated by sediment toxicity to amphipods, there was a high confidence of potential effects to benthic receptors (Figure 7-8). Measures of exposure indicated a medium confidence of high exposure (Figure 7-8). Negligible exposure was suggested by AVS concentrations that were much higher than divalent metal concentrations. , High exposure was suggested by enriched levels of As, Cr, Pb, Hg, and Ag in

10). These data were interpreted to indicate that there was medium confidence of potential effect and elevated exposure to salt marsh receptors in the Sullivan Point AOC.

Issues associated with the measures of salt marsh ecology at Sullivan Point included the significance of the relatively high abundance of other vascular plants and reduced abundance of amphipods and other animal taxa. The relatively high abundance of other vascular plants suggests that the marsh was stressed because it appeared that the ability of dominant *Spartina* grasses to outcompete other plants was diminished. However, no impact was suggested from data that showed the marsh had relatively high biomass, height, and percent cover of *Spartina* and that high numbers of species were present in the SAT zone. Also indicating stress was the reduced vigor of the plants along the seaward edge of the marsh. This could lead to erosion of the peat, loss of habitat, and the release of contaminants bound to the peat.

7.1.6.6 Sullivan Point Summary

The Sullivan Point area had indications of negligible risk to pelagic receptors, low risk to epibenthic and eelgrass receptors, and intermediate risk to benthic and salt marsh receptors (Table 7-5). For the pelagic assessment endpoint, the confidence in the conclusion of negligible risk was based on the agreement between negligible estuarine surface-water concentrations and the absence of surface-water toxicity. There was a medium level of confidence associated with the finding of low risk to epibenthic and eelgrass receptors, medium confidence of intermediate risk to salt marsh receptors, and a high level of confidence in the conclusion of intermediate risk to benthic receptors (Table 7-5). The medium confidence of negligible risk to pelagic receptors in the Sullivan Point AOC was due to the agreement between negligible estuarine surface-water concentrations and the absence of surface water toxicity. High confidence in intermediate risk to benthic receptors in the Sullivan Point AOC was due to the concordance between exposure (exceeding ER-M threshold and predicted pore water toxicity for DDT) and effect (sediment toxicity) measures.

7.1.7 DRMO Storage Yard

Characterized by strong currents and the lack of sediment accretion, the DRMO Storage Yard area AOC is located directly adjacent to the portion of the river that has some of the highest current velocities measured in the whole estuary. Shored up with riprap and large boulders, the shoreline along the Storage Yard is being eroded by current and wave action. The bottom of the river near the Storage Yard consists of boulders, cobbles, gravel and sandy gravel and no appreciable accumulation of sediment is present.

7.1.7.1 DRMO Storage Yard Pelagic

For the pelagic assessment endpoint, monthly monitoring of phytoplankton biomass and the absence of toxicity to *Arbacia* fertilization indicated no effect to pelagic receptors. Because there was agreement of no effect between the measures of phytoplankton biomass and *Arbacia* toxicity, a medium level of confidence was assumed. The only pelagic exposure measure was the estuarine surface water concentrations that were well below ambient WQC levels. The exposure

7.1.8 Dry Dock

The most heavily industrialized portion of the whole estuary is located within the Dry Dock AOC. Three operating dry docks, numerous tug boats and support vessels, and the major industrial area of the Shipyard are located and operated within the Dry Dock AOC. Open to high currents from the main channel and the connection between the main channel and back channel, most of the area is well flushed. Eddies and restricted flow of river water during certain tidal conditions results in the accumulation of deposits of sandy-muds. The deposits are found in small areas adjacent to the berthing areas near the dry docks and along the ledge that connects Wattlebury and Badger Islands. There were two established eelgrass beds within the Dry Dock AOC, one was located near Wattlebury Island (station 17) and the other was located between dry docks 1 and 2 (station 12.5). There was no salt marsh habitat within the Dry Dock AOC.

7.1.8.1 Dry Dock Pelagic

Effects to pelagic receptors in the Dry Dock AOC were assessed by the measurement of toxicity to *Arbacia* fertilization at four stations and exposure was assessed by measurements of estuarine surface water concentrations at five stations. No toxicity was observed and surface water samples near the Dry Docks were relatively higher than other areas in the lower estuary, but the concentrations were well below ambient WQC. Because no toxicity was observed and estuarine surface water concentrations were well below ambient WQC it was concluded that there was a medium level of confidence of no effect and negligible exposure to pelagic receptors in the Dry Dock AOC (Figure 7-13).

Issues associated with pelagic assessment endpoint for the Dry Dock area were the lack of caged mussel data for this site and the possible sources of contamination to the area. Caged mussels deployed in the Dry Dock AOC during Phase I were lost because the cages interfered with tug boat operations. Since the surface waters near the Dry Docks are also influenced by discharge from the Portsmouth sewage treatment plant, there was uncertainty about what contribution of contamination could be coming from the Shipyard. In addition, the holding times for the *Arbacia* toxicity test were exceeded and there was no information about current discharges from the shipyard into the Dry Dock AOC.

7.1.8.2 Dry Dock Epibenthic

The measures of epibenthic effects showed potential effects indicated by mussel density[8] and reduced fucoid alga biomass at one of the two fucoid algae monitoring stations. No effect was indicated by the greater than average densities of lobsters and similar to reference levels of mussel condition index and mussel shell length. High exposure was indicated by mussel residues which exceeded critical values for PHEN, were near critical values for Pb, and exceeded warning levels for Ag, Cu, Cr, Ni, and Hg. Elevated exposure was indicated by the significantly higher than reference concentrations of PAHs and the highest concentrations of Hg measured in lobster tissues. Low exposure was indicated by the higher than average concentrations of As in fucoid alga tissue, and negligible exposure was indicated by the estuarine surface water concentrations

---

[8] Because mussels were absent from station 13, a potential effect to mussel density was assumed

richness, and evenness were measured at the other stations and only moderate toxicity (mortality = 90%) was detected at one other station (station 17), the spatial extent of the impacted area may be limited. Since higher exposure levels were measured at other stations where no effects were observed, it is uncertain what caused the toxicity and benthic anomaly at station 13. There was also uncertainty about the seasonal variation of AVS levels, the large variability in the benthic community among stations in the Dry Dock AOC, and the potential long term impact to benthic receptors. Confidence was increased by the relatively large sample size for analytical chemistry (n = 33 for sediments) and benthic community and toxicity analyses (n = 16 — 4 stations × 4 reps).

7.1.8.4 Dry Dock Eelgrass

For measures of effects to eelgrass in the Dry Dock AOC, potential effects were suggested by low shoot density, the lack of reproductive shoots. No effect was indicated by the similarity to reference of the leaf : shoot ratio, leaf and root/rhizome morphology[11], and the contiguous, healthy growth of the eelgrass beds studied. For measures of exposure, high exposure was indicated by bulk-sediment concentrations which exceeded ER-L levels for Pb and Hg, and elevated exposure was suggested by leaf residues above the ETC for Pb and Cu and root/rhizome residues above the ETC for Pb, Hg, Cr, and Cu. From these data it was concluded that there was medium confidence of no effect and elevated exposure to eelgrass in the Dry Dock AOC (Figure 7-16).

Issues related to the eelgrass assessment endpoint for the Dry Dock AOC were identified. There were two contiguous eelgrass beds with healthy growth present (at stations 12.5 and 17) within the AOC, the plants appeared to be healthy but there were relatively high concentrations of contaminants measured in the leaf and root/rhizome tissues. Data on eelgrass morphology, density of reproductive shoots and vegetative shoots, the number of leaves per shoot were not collected for station 12.5, so there was uncertainty about possible differences between the two eelgrass beds. The area of suitable eelgrass habitat in the Dry Dock AOC was not quantified. Because benchmarks for eelgrass are not available it is uncertain what effect contaminant exposure and accumulation had on the plants.

7.1.8.5 Dry Dock Summary

For the Dry Dock AOC negligible risk was determined for pelagic receptors, low risk was indicated for the epibenthic and eelgrass assessment endpoints, and intermediate risk was concluded for benthic receptors. There was a medium level of confidence in the conclusion of negligible risk determined for pelagic receptors and in the low risk indicated for epibenthic and eelgrass assessment endpoints. For the benthic receptors there was a high confidence level in the conclusion that the magnitude of risk was intermediate (Table 7-7). The high confidence of intermediate risk to benthic receptors in the Dry Dock AOC was due to the concordance between exposure (bulk-sediment chemistry and predicted pore water toxicity) and effect (sediment toxicity and low invertebrate density) measures. Because there were no salt marsh habitats in the Dry Dock AOC, exposure and effects to salt marsh receptors were not assessed for this AOC.

---

[11] No data on leaf and root/rhizome morphology were available for the eelgrass bed located at station 12.5.

whether the high tissue concentrations measured in the mussels had any impact on the mussel population.

7.1.9.3 Back Channel Benthic

For measures of effect to the benthic community in the Back Channel AOC, probable effects were indicated by amphipod toxicity and no effect was indicated by the similarity of invertebrate abundance, species richness, and species evenness to reference conditions. For exposure measures, adverse exposure was indicated by pore water concentrations of PHEN, ANTH, FLUORAN, PYRENE, and DDT that were greater than EqP-predicted porewater toxic units and bulk-sediment concentrations that where higher than the ER-M for PHEN. Elevated exposure was suggested from the enriched levels of Pb, Hg, and Cd in the sediment. From these data it was concluded that there was high confidence in potential effects and medium confidence in high exposure to benthic receptors in the Back Channel AOC (Figure 7-19).

Since the Back Channel area is relatively large with such diverse bottom characteristics, a major issue was that the area was not well characterized. Benthic effects were assessed at only one station (station 18 — four replicates) and there were no data on AVS-SEM concentrations in the Back Channel AOC. There was good correspondence between indications of the bioavailability of PAHs and DDT estimated by EqP and the finding of toxicity at station 18. As suggested by the toxicity to amphipods, ampeliscids were also absent from the benthic community sampled at station 18. Because the measures of benthic community structure did not indicate an effect, there may of been a shift to more pollution tolerant species at the site. Capitella capitata and oligochaetes dominated the benthic community at station 18. Both of these species (taxa) have been identified as pollution-tolerant (Levin 1984).

7.1.9.4 Back Channel Eelgrass

Measures of effects on eelgrass in the Back Channel AOC showed potential effects indicated by reduced root/rhizome biomass, the absence of reproductive shoots, and the small, limited distribution of the eelgrass bed. No effect was indicated by the similarity to reference of measures of shoot density, leaf : shoot ratio, and leaf morphology. Exposure measures showed high exposure from bulk chemistry concentrations that exceeded ER-L levels for Pb, Cd, and Cu and elevated exposure from leaf and root residues that exceeded ETC levels for Pb, Hg, Cr and Cu. From these data it was concluded that there was medium confidence of potential effects and elevated exposure to eelgrass in the Back Channel AOC (Figure 7-20).

The major issue for the eelgrass assessment endpoint in the Back Channel was that the eelgrass bed sampled in the Back Channel was small and limited in distribution. The limited distribution may be due to chemical exposure, but suitable substrate and other physical factors may also limit the eelgrass distribution. The area of suitable eelgrass habitat in the Back Channel was not quantified. There is also uncertainty why reproductive shoots were absent when leaf biomass and shoot density appeared to be normal. The reduced root/rhizome biomass was correlated with high Pb exposure, but because benchmarks for eelgrass effects are not available it is uncertain what effect contaminant exposure and accumulation has on the plants.

7.1.10.1 Jamaica Cove Pelagic

The measure of effects to Pelagic receptors in Jamaica Cove showed no effects. No impact to deployed mussel scope-for-growth (SFG) or toxicity to sea urchin fertilization was detected. Exposure measures showed high exposure from seep water samples that exceeded ambient WQC for Cu, Ni, Zn, and Pb, and elevated exposure from the Phase II deployed mussels which accumulated Pb concentrations at significantly higher levels than predeployed mussels. Negligible exposure was indicated by estuarine surface water concentrations that were well below ambient WQC and the lack of accumulation of contaminants in the mussels deployed during Phase I. From these data it was concluded that there was medium confidence of no effects and low exposure to pelagic receptors in Jamaica Cove (Figure 7-22).

Issues were identified for pelagic measurement endpoints in Jamaica Cove. Although mussels deployed during Phase II resulted in the significant accumulation of Pb above predeployment levels, there was no difference among the Pb accumulated in mussels deployed in Jamaica Cove, Clark Cove, and the reference location in Pepperrell Cove. This suggests that there was not a gradient of Pb exposure levels during the deployment period and that the accumulation may be more of a function of the depositional characteristics of the coves rather than indications of potential sources of Pb. Because the Phase I seep samples from Jamaica Cove did not meet data quality objectives,[12] the data were not used in the risk characterization. Uncertainty was also increased because holding times for the *Arbacia* toxicity test were exceeded, there were differences in the length of caged mussel deployments between Phase I and II, and there was a lack of knowledge of how quickly seeps would be diluted upon entering Jamaica Cove.[13] The relatively high number of seep (n = 20) and estuarine surface water samples (n = 13) analyzed for chemical constituents increased the confidence in the exposure assessment.

7.1.10.2 Jamaica Cove Epibenthic

Measures of effects to epibenthic receptors in Jamaica Cove showed no effect. Measures of fucoid biomass, mussel density, mussel shell length, and mussel condition index were greater than or equal to reference conditions. High exposure was suggested from seep waters, which exceeded ambient WQC for Pb, Cu, Ni, and Zn. Mussel tissue residues that exceeded the critical value for Pb and warning levels for Pb, Cu, Zn, Ni, and Hg also suggested high exposure. Elevated exposure was indicated by lobster tissue concentrations that were significantly higher than reference concentrations for PAHs and Hg. Negligible exposure was suggested by the very low concentrations measured in estuarine surface water samples and by fucoid alga residues which were below average. From these data it was concluded that there was medium confidence of no effect and elevated exposure to epibenthic receptors in Jamaica Cove (Figure 7-23).

Issues associated with the epibenthic measures for Jamaica Cove were uncertainties about the effects of exposure from the seeps and whether high residue concentrations in the mussels and

---

[12] The Phase I seep samples were contaminated during sampling by the entrainment of sediment particles in the water samples.

[13] Data obtained during the Gulf Challenger cruises suggests that during ebb tide conditions the seeps were diluted to sub-ppb levels in less than 10 meters from the discharge point (Cullen and Arimoto 1995).

Hg, and Cu. From these data it was concluded that there was medium confidence of no effect and elevated exposure to eelgrass in Jamaica Cove (Figure 7-25).

An issue for the eelgrass assessment endpoint in Jamaica Cove was the physical size of the eelgrass bed. Although the eelgrass bed in Jamaica Cove was small, it appeared to be healthy and there was uncertainty about what may be limiting its distribution. The distribution may be limited due to chemical exposure, but suitable substrate and other physical factors may also limit eelgrass distribution. The area of suitable eelgrass habitat in Jamaica Cove was not quantified. Because sediment-Pb concentrations were higher in other areas, the reduced root/rhizome length did not appear to be correlated with high Pb exposure. Since benchmarks for eelgrass effects are not available, it is uncertain what effect contaminant exposure and accumulation may have on the plants.

7.1.10.5 Jamaica Cove Salt Marsh

Other than reduced levels of amphipod abundance, all the effect measures for salt marsh receptors in Jamaica Cove showed no effect. Measures of *Spartina* cover, *Spartina* morphology, mollusc abundance, number of animal taxa, other vascular plant cover, and the ratio of live:dead *Littorina* were all similar to measures made in the reference marshes. Exposure measures showed high exposure from bulk-sediment concentrations that were above ER-L thresholds for Pb and Zn, and negligible exposure from the relatively low concentrations of metals in *Spartina* tissues. These data suggested that there was medium confidence of no effect and elevated exposure to salt marsh receptors in Jamaica Cove (Figure 7-26).

In addition to the lack of effects benchmarks to evaluate the effects of exposure to salt marsh ecology another issue was what may of caused the reduced number of amphipods present in the Jamaica Island marsh. The Jamaica Island marsh had poorly developed peat and lacked a high marsh community (probably because of the steep road embankment directly behind the marsh). Nevertheless, the effect measures showed that the plants in the low marsh were healthy and would probably continue to promote the accumulation of peat and prevent the erosion of sediments and the release of contaminants associated with the sediments.

7.1.10.6 Jamaica Cove Summary

For Jamaica Cove the weight of evidence showed negligible risk to pelagic and benthic receptors, and low risk to epibenthic, eelgrass, and salt marsh assessment endpoints (Table 7-9). Based on the weights assigned to the measures of exposure and effects, the confidence in the magnitude of risk were medium for pelagic, epibenthic, eelgrass, and salt marsh assessment endpoints and high for the benthic assessment endpoint (Table 7-9). There was high confidence of negligible risk to benthic receptors in Jamaica Cove due to the concordance between highly weighted effect measures.

7.1.11 Portsmouth Harbor

The exposure and effects measures that were applicable to Portsmouth Harbor as a whole were assessed to determine the weight of evidence of risk to pelagic, epibenthic, and avian assessment endpoints. Included within the Portsmouth Harbor focus area were data from the

A major issue was how to correctly interpret the winter flounder data. Even though the sampling program was conducted to target winter flounders that are year-round residents of the estuary, the migratory nature of winter flounder resulted in a high degree of uncertainty of where the flounder accumulated contaminants and how the spleen pathology was initiated. There appeared to be a significant correlation between the incidence of SMA and PAH exposure. The similarity between the Portsmouth Harbor and Gulf of Maine populations could mean that migrants between both areas intermingle and that elevated exposures to winter flounder occur on a regional scale. Uncertainty was also increased by the limited numbers of samples and the compositing of samples among individuals that was required to obtain enough material for accurate chemical analyses.

7.1.11.2 Portsmouth Harbor Epibenthic

Other than the potential effect suggested by reduced condition indices in mussels sampled from Portsmouth Harbor, the measures of effects to epibenthic receptors in Portsmouth Harbor showed no effect. Measures of lobster density, mussel density, and mussel lengths were greater than or equal to measures made in reference areas (upper estuary and York River). For exposure measures, high exposure was indicated by mussel tissue concentrations that exceeded critical values for Pb and some PAHs. Elevated exposure was indicated by juvenile lobster flesh tissues which had significantly higher than reference concentrations of PAHs and Hg. Juvenile lobster hepatopancreas tissues were significantly higher than reference for PAHs, adult lobsters (both legal and sublegal) had significantly higher concentrations of PAHs in hepatopancreas tissues, and adult lobsters (both legal and sublegal) had higher than reference (but not significant) concentrations of Hg in flesh tissues. Negligible exposure was indicated by the very low estuarine surface water concentrations. From these data it was concluded that there was medium confidence of no effect and elevated exposure to epibenthic receptors in Portsmouth Harbor (Figure 7-28).

Issues associated with the epibenthic measures for Portsmouth Harbor were identified. There was uncertainty about the sources of exposure in the Harbor and the migratory nature of adult lobsters in the estuary. There were no effect benchmarks for tissue residues so the ability to determine whether high residue concentrations in the mussels and lobsters could cause any effects was limited. There was also a high degree of uncertainty about the source of exposure and route of uptake of the contaminants measured in lobster tissues. The relatively high number of mussel, lobster, and surface water samples analyzed for chemical constituents increased the confidence in the exposure assessment.

7.1.11.3 Portsmouth Harbor Avian

The exposure assessment for avian receptors in the Portsmouth Harbor focus area showed negligible exposure to avian receptors because the calculated hazard index (HI) was less than 2 for all dietary pathways. The dietary uptake of contaminants modeled for herbivores resulted in a HI < 1 for Canada goose consumption of eelgrass leaves, eelgrass roots, and *Spartina* leaves, and a HI = 1.6 for Canada goose consumption of fucoid algae. The dietary uptake of contaminants modeled for omnivores resulted in a HI < 1 for black duck consumption of eelgrass leaves, eelgrass roots, and *Spartina* leaves, a HI = 1.6 for black duck consumption of fucoid algae, and a HI = 1.7 for black duck consumption of blue mussels. The dietary uptake of contaminants

where

R_i = magnitude of risk

C_i = confidence of conclusion, and

WM_i = weight applied to evaluate risk from media.

The magnitude of risk (R_i) and confidence in conclusion (C_i) from the summary tables (Tables 7-4 to 7-10) were assigned numeric values (Table 7-11) to calculate the magnitude of risk from medium (R_M) and confidence in conclusions from medium (C_M). The lookup cut off values (Table 7-11) were intervals determined to interpret the results of the calculations. For example the magnitude of risk/confidence in conclusions (R_i/C_i) determined for Clark Cove (Table 7-4) were pelagic—Low/Medium (1/2), epibenthic—Low/Medium (1/2), benthic—Low/High (1/3), eelgrass—Intermediate/Low (2/1), and saltmarsh—Low/Medium(1/2) resulted in calculations of magnitude of risk from medium/confidence in conclusions from medium (Table 7-13, R_M/C_M) for water—Low/Medium (1.2/2) and sediment—Low/High (1.2/2.4).

The weighing scheme was used to calculate the entries for Table 7-13. For areas of concern where all endpoints were not assessed (*e.g.* The Dry Dock area was not assessed for impacts to salt marsh receptors, and the DRMO was not assessed for impacts to benthic, eelgrass, and salt marsh receptors) the missing assessment endpoint(s) was(were) excluded from the calculation. For the Portsmouth Harbor focus area the surface water medium was evaluated for the pelagic and epibenthic assessment endpoints and the biota medium was evaluated as dietary exposure to avian receptors. Footnotes were provided to document significant findings and explain the use of professional judgment.

The characterization of ecological risk associated with environmental media by AOC (Table 7-14) showed that there was high confidence of intermediate risk from sediment exposure the Back Channel, Sullivan Point, and Dry Dock areas of concern. High confidence of low risk was evident for sediment exposure in Clark Cove and Jamaica Cove. There was medium confidence of low risk for surface water exposures in Clark Cove, Sullivan Point, the Dry Dock the Back Channel, Jamaica Cove, and the Portsmouth Harbor focus area. There was also medium confidence in negligible risk from surface water exposure at the DRMO Storage Yard, and medium confidence of negligible risk from dietary exposure to avian receptors in Portsmouth Harbor.

Uncertainty associated with the characterization included the following factors:

- evidence of bioaccumulation in all of the AOCs is probably related to surface water exposures;
- evidence of bioaccumulation in juvenile lobsters measured in Clark Cove, Jamaica Cove, Sullivan Point, and the Dry Dock AOC is probably related to sediment exposure levels; and
- in areas where very fine-grained sediments are present (*e.g.* Clark Cove), sediment resuspension may be influencing surface water risks.

(sensitivity of the measure to stressor responses). For example, one sampling event is extremely useful for determining the synoptic conditions associated with the site, but it provides no information on temporal changes. Only idealistic ecorisk assessments with unlimited time and money could possibly score high for all of the study design attributes (how well the measure represented site conditions, was sensitive to stressor responses, represented spatial differences, represented temporal or seasonal variations, provided quantitative data, and used standard and acceptable methodology (Menzie *et al.* 1996)). However, improvements in the design and execution of the studies conducted to obtain the measurement data could result in more highly weighted measures for inferring risk. For example, periodic sampling conducted as part of a monitoring program would help increase the weight of many of the measures.

Indicating moderate to low levels of uncertainty in the measurement data, most of the measures had weights of high data quality, medium strength of relationship, and medium study design. This was because the studies conducted for the risk assessment were site-specific, were directed at specific ecological components and receptors within the study area, followed standardized sampling and analysis procedures, complied with appropriate quality control and quality assurance procedures, and provided measurements that were applicable to the assessment endpoint.

Even though the endpoint weights were assigned after studies for the risk assessment had been completed, the endpoint weighting exercise provided a basis to reach a consensus and formalize the collective professional judgment of the Ecorisk Team for interpreting the results. This helped focus the characterization of risk on the evidence provided by the data. By assigning endpoint weights at the beginning of the risk assessment (during problem formulation), better use could be made of resources available to conduct the risk assessment and clearer descriptions of risk could be developed.

7.1.13.2 Evidence of Risk

Effect measures were evaluated to determine whether the relative weight added to the conclusion that effects to the endpoint were or were not evident. Exposure measures were evaluated relative to the conclusion that exposure would or would not cause an effect. In this sense exposure measures for which effects benchmarks were available (e.g. surface water concentrations) could be much better evaluated than exposure measures that did not have effects benchmarks (e.g. tissue residues). When benchmarks were not available, it was necessary to compare the results obtained from the AOC to data obtained from reference areas. Reference areas were used to evaluate effects relative to "pristine" areas (Isles of Shoals, Brave Boat Harbor) as well as other localized sources (upper estuary, York Harbor, lower Piscataqua River, Spruce Creek, etc). Reference data used for comparison purposes carried the same relative weight of the measure being evaluated. The appropriateness of the reference data was evaluated as part of the "study design" contribution to the measure's endpoint weight.

The magnitude of risk was determined by the combination of conclusions from the exposure and effects data (see Table 7-2). Each measure was weighted by its ability to assess risk to the assessment endpoint. For example, in the assessment of risk to epibenthic endpoint in Clark Cove (Figure 7-2), data from the mussel and seep samples showed "high exposure", but data from surface water and fuciod samples showed "negligible exposure" and data from the lobster samples showed "elevated exposure" . This resulted in a conclusion of "elevated exposure". However, the

evidence and developing a consensus in the conclusions derived, personal or professional biases were eliminated as much as possible.

One of the objectives of the risk assessment is to clearly communicate the risks detected. Conveying the actual measurement outcomes in summaries that highlighted the salient results (Appendix IX) helped ensure that information that may be important in management decisions was not lost. A critical aspect of risk communication involves identifying uncertainties inherent in measurement activities, data interpretation, and the risk assessment process itself. Confidence was established by assigning weights to the measurement endpoints and evaluating the degree of preponderance, magnitude, and concurrence among the endpoint responses. Specifying the issues and uncertainty associated with individual measures and their interpretation and independently evaluating exposure and effect information enhanced the level of confidence in the final conclusions. Further, during each level of progressively summarizing the information, the confidence in the conclusions reported was reevaluated. Because each consolidation step involved progressively greater amounts of information and additional synthesis, the confidence in the final conclusions generally increased as the process proceeded (Munns et al. in prep).

When the risk was characterized risk for each of the AOCs, care was taken to qualify the conclusions by narrating the supporting rational. The narrative included issues that reinforced the confidence and identified major sources of uncertainty that were associated the conclusions advanced. This information is valuable to risk managers and other stake holders because it helps make the process of characterizing risk easier to understand, more transparent, and hopefully more widely acceptable. Even though some might not agree with the conclusions, the process assures that they can clearly see how the conclusions were derived.

7.1.13.3 Risks From Environmental Media

Assuming that the behavior of contaminants released in the estuary would follow the hypothesized exposure pathways, the assessment endpoints were expected to respond differently to different routes of exposure. While it is very difficult to completely separate exposure from water, sediment, and food, it was assumed that the pelagic and benthic assessment endpoints would most likely be more affected by surface water and sediment exposure, respectively. Many infauna species filter the overlying surface water (e.g. clams, suspension feeding polychaetes, etc.) and some of the pelagic receptors are actually demersal foragers (e.g. winter flounder) that spend a large portion of their life in close contact with the sediment. However, it was assumed that the predominant route of exposure was from the sediment for the benthic assessment endpoint and from the water for the pelagic assessment endpoint. For the epibenthic, eelgrass, and salt marsh assessment endpoints, the assumption was that exposures from water or sediment were equally possible. For example, the accumulation of contaminants by mussels is a measure of exposure to the epibenthic community that could be a result of either surface water or (suspended) sediment exposure. For avian receptors the major route of exposure was dietary (Figure 7-30).

In the analysis, the conclusion of risk from the exposure media (water, sediment, diet) was based on the level of risk determined for each assessment endpoint, weighted by the predominant route of exposure (Table 7-13). Low risk for the media would be concluded if any of the assessment endpoints for which the media was the predominant route of exposure indicated low risk. In the case for Clark Cove (Table 7-4) the magnitude of risk was "Low" for the pelagic, benthic, epibenthic, and salt marsh assessment endpoints and "Intermediate" for eelgrass. In order

Table 7-1. Scheme used to interpret outcomes of measurement activities.

Type of Measure	Degree of Response	Interpretation	Numerical Value (M_i)
Exposure	≤ reference condition or below conservative benchmark concentration	negligible exposure	0
	> qualitative screening level	low exposure	1
	statistically > reference condition	elevated exposure	2
	> a conservative benchmark concentration	high exposure	3
	> a nonconservative benchmark concentration	adverse exposure	4
Effect	similar to reference condition or below ecologically-relevant threshold	no effect	0
	worse than reference condition, but not statistically different[1]	potential effect	1
	statistically worse than reference or control condition[1]	probable effect	2

[1] The data from the AOC are evaluated to determine if there is a problem relative to the reference area. In some cases lower than reference is desirable (*e.g.* percent cover of other vascular plants in salt marsh lower than reference), while in other cases higher than reference is desirable (*e.g.* fucoid algae biomass higher than reference).

Table 7-2. Interpretation of exposure and effect evidence in determining risk.

	Evidence of Exposure				
Evidence of Effect	NEGLIGIBLE	LOW	ELEVATED	HIGH	ADVERSE
NO	Negligible	Negligible	Low	Low	Intermediate
POTENTIAL	Negligible	Low	Intermediate	Intermediate	High
PROBABLE	Low	Low	Intermediate	High	High

Table 1-3. Values of effect (EF_i) and exposure (EX_i) versus endpoint weights (EW_i). used for constructing scatter plots (A) and example calculation (B).

A. Values of effect (EF_i) and exposure (EX_i) versus endpoint weights (EW_i).

Effects Measures (x)	Exposure Measures (x)	Endpoint Weights (y)
Outcome = (EF_i)	Outcome = (EX_i)	Endpoint Weight = (EW_i)
No = 1	Negligible = 1	Low = 1
Potential = 2	Low = 2	Medium = 2
Probable = 3	Elevated = 3	High = 3
	High = 4	
	Adverse = 5	

B. An example calculation for measures of exposure and effects to the Pelagic Assessment Endpoint in Clark Cove.

Raw Data entered from summary table (Appendix IX.1)

	Effect			Exposure	
name	EF_i	EW_i	name	EX_i	EW_i
Phytoplankton Biomass	1	1	Surface Water	1	2
Deployed Mussel SFG	1	2	Deployed Mussel I	1.1	2
Arbacia Toxicity	3	2	Deployed Mussel II	3	2
			Seep	4	1
Effect			Exposure		
Plotted Values offset by -.5 in x and y direction					
name	EF_i	EW_i	name	EX_i	EW_i
Phytoplankton Biomass	0.5	0.5	Surface Water	0.5	1.5
Deployed Mussel SFG	0.5	1.5	Deployed Mussel I	0.6	1.5
Arbacia Toxicity	2.5	1.5	Deployed Mussel II	2.5	1.5
			Seep	3.5	0.5

Weighted Average Calculations to Determine Centroid Location

	Weighted Effects			Weighted Exposure	
name	EW_i	EF_i *EW_i		EW_i	EX_i *EW_i
Phytoplankton Biomass	0.5	0.25	Surface Water	1.5	0.75
Deployed Mussel SFG	1.5	0.75	Deployed Mussel I	1.5	0.9
Arbacia Toxicity	1.5	3.75	Deployed Mussel II	1.5	3.75
			Seep	0.5	1.75
	ΣEW_i	ΣEF_i *EW_i		ΣEW_i	ΣEX_i *EW_i
sum	3.5	4.75		5	7.15
(ΣEF_i *EW_i)/ΣEW_i sum(effects*ew)/sum(ew)		1.3571	(ΣEX_i *EW_i)/ΣEW_i sum(exp*ew)/sum(ew)		1.43
	effect	ew		exp	ew
	avg(W_x)	avg(y)		avg(W_x)	avg(y)
Plotted Centroid	1.35	1.1667		1.43	1.25

Table 7-6. Summary of risk to assessment endpoints in the DRMO Storage Yard area of concern.

Assessment Endpoint	Evidence of Effect [1]	Evidence of Exposure [1]	Magnitude of Risk [1]	Confidence in Conclusions [1]
Pelagic	No/M	Negligible/M	Negligible	Medium
Epibenthic	No/M	Elevated/M	Low	Medium
Benthic [2]				
Eelgrass [2]				
Salt Marsh [2]				

[1] See Table 7-4 for description.
[2] No sediment, eelgrass, or salt marsh habitat in this area of concern.

Table 7-7. Summary of risk to assessment endpoints in the Dry Dock area of concern.

Assessment Endpoint	Evidence of Effect [1]	Evidence of Exposure [1]	Magnitude of Risk [1]	Confidence in Conclusions [1]
Pelagic	No/M	Negligible/M	Negligible	Medium
Epibenthic	No/M	Elevated/M	Low	Medium
Benthic	Potential/H	High/M	Intermediate	High[3]
Eelgrass	No/M	Elevated/M	Low	Medium
Salt Marsh [2]				

[1] See Table 7-4 for description.
[2] No salt marsh habitat in this area of concern.
[3] High confidence due to the concordance between exposure (bulk-sediment chemistry and predicted pore water toxicity) and effect (sediment toxicity and low invertebrate density) measures.

Table 7-10. Summary of evidence of risk to assessment endpoints for Portsmouth Harbor focus area.

Assessment Endpoint	Evidence of Effect [1]	Evidence of Exposure [1]	Magnitude of Risk [1]	Confidence in Conclusions [1]
Pelagic	No/M	Elevated/M	Low	Medium
Epibenthic	No/M	Elevated/M	Low	Medium
Benthic [2]				
Eelgrass [2]				
Salt Marsh [2]				
Avian	Not Evaluated	Negligible/M	Not Evaluated [3]	Medium [4]

[1] See Table 7-4 for description.
[2] Benthic, eelgrass, salt marsh endpoints not evaluated for Portsmouth Harbor focus area.
[3] With the lack of effects information, the most conservative estimate of risk is Low.
[4] Pertains to the confidence of exposure measures only.

Table 7-13. Characterization of ecological risks associated with environmental media at Portsmouth Naval Shipyard by Area of Concern.

Area of Concern	Environmental Medium	Magnitude of Risk From Medium	Confidence In Conclusions
Clark Cove	Surface Water[1] Sediment[3]	Low[2] Low	Medium High
Sullivan Point	Surface Water Sediment[3]	Low Intermediate	Medium High
DRMO[4]	Surface Water[1]	Negligible	Medium
Dry Docks[5]	Surface Water[1] Sediment[3]	Low Intermediate	Medium High
Back Channel	Surface Water[1] Sediment	Low Intermediate	Medium High
Jamaica Cove	Surface Water[1] Sediment[3]	Low Low	Medium High
Portsmouth Harbor[6]	Surface Water[1] Biota[7]	Low Negligible	Medium Medium

[1] Evidence of bioaccumulation in mussels is probably related to surface water exposure.

[2] Sediment resuspension may be influencing surface water risks.

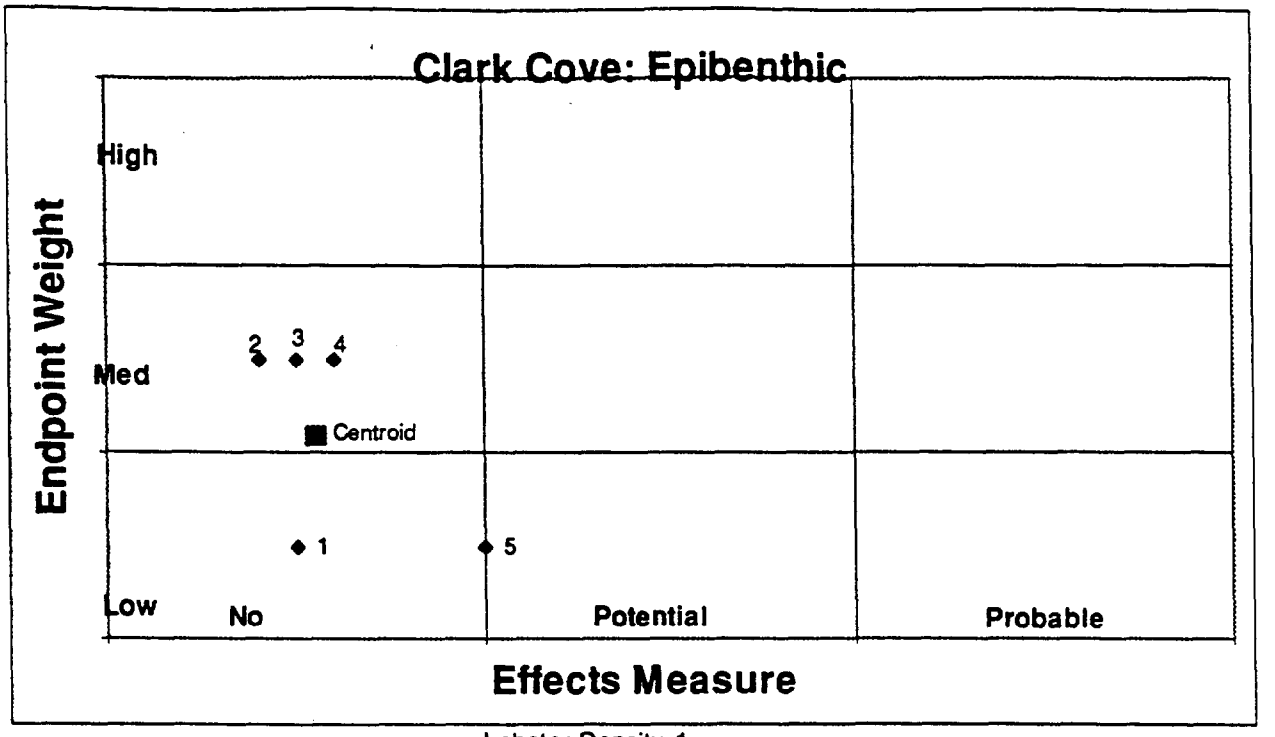
[3] Evidence of bioaccumulation in juvenile lobsters may be related to sediment exposure.

[4] No sedimentary, eelgrass, or salt marsh habitat at DRMO.

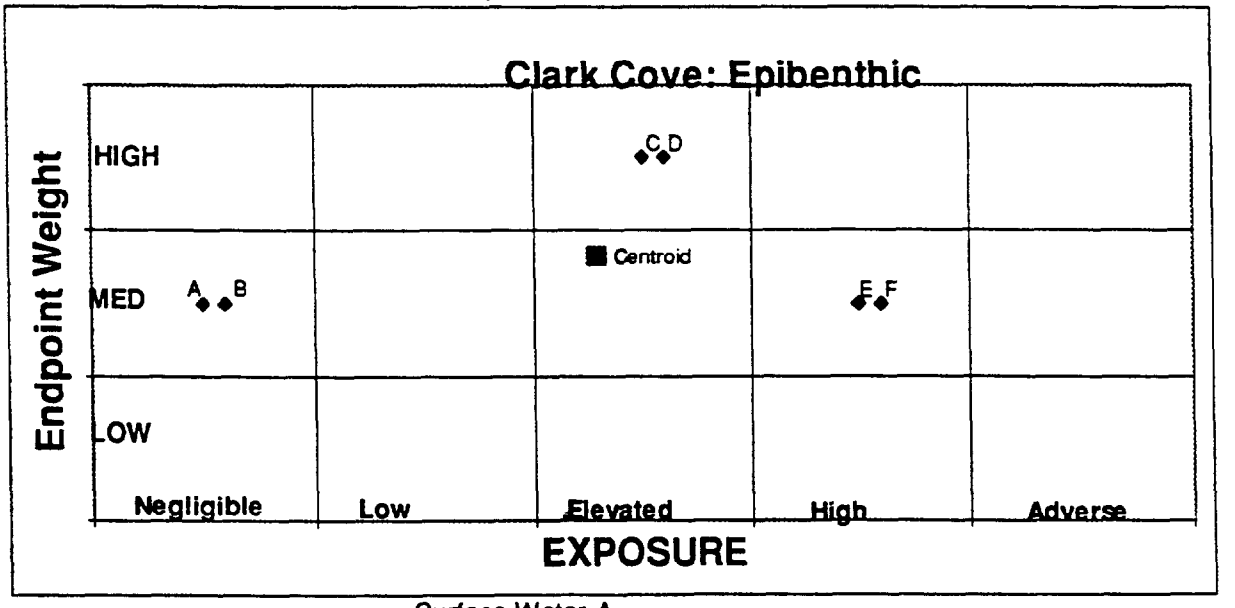
[5] No salt marsh habitat in Dry Dock area of concern.

[6] Only pelagic, epibenthic, and avian assessment endpoints were evaluated for Portsmouth Harbor focus area.

[7] Biota evaluated as dietary exposure to avian receptors.

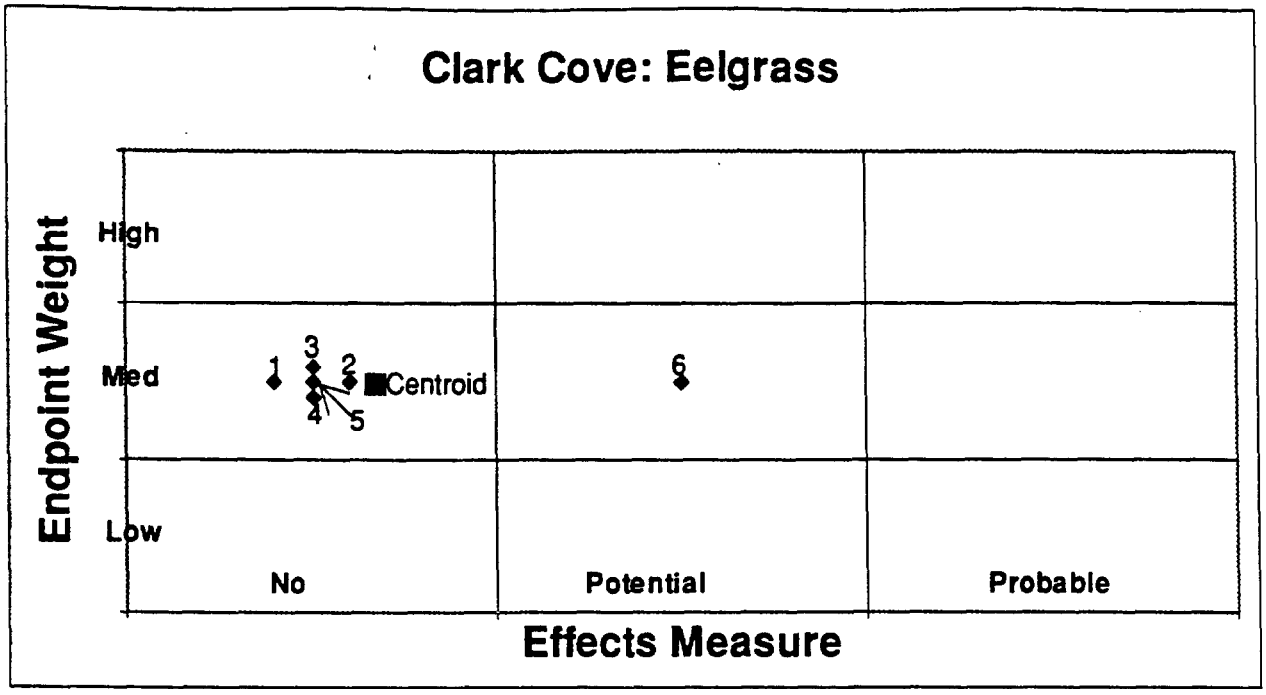


Lobster Density 1
Mussel Density 2
Mussel Length 3 No/M
Mussel Condition Index 4
Fucoid Biomass 5



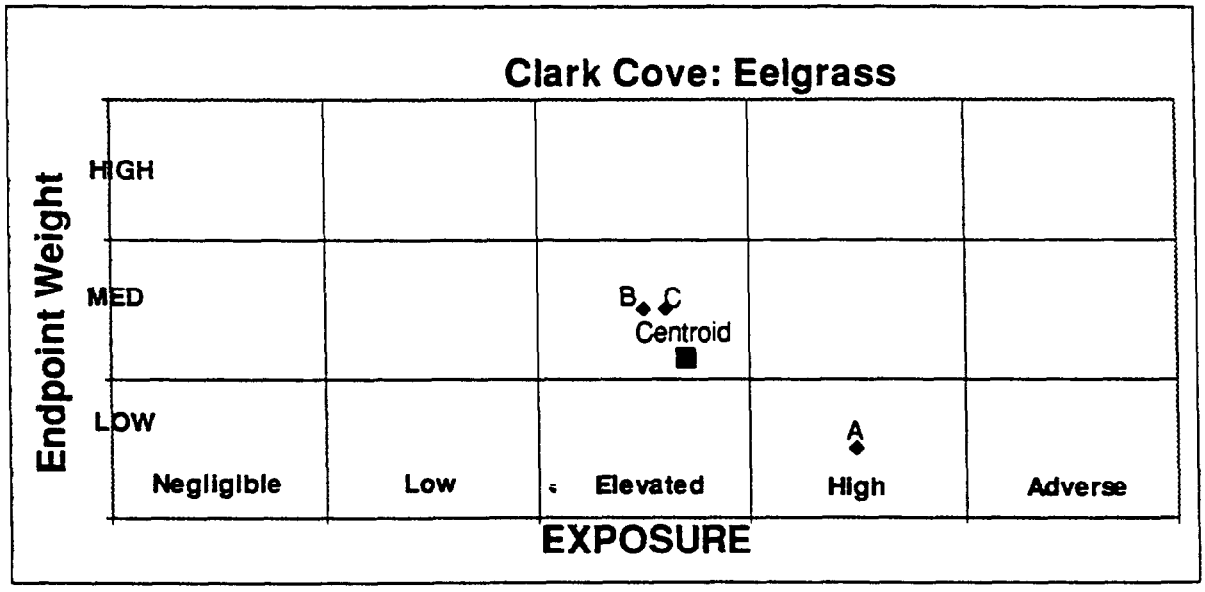
Surface Water A
Fucoid Res B
Juvenile Lobster Hepato Residues C Elevated/M
Juvenile Lobster Tail+Claw Residues D
Seeps E
Mussel Residues F

Figure 7-2. Epibenthic assessment data for Clark Cove and weight of evidence from exposure and effects measures.



Root Morphology 1
Leaf Morphology 2
Shoot Density 3
Reproductive Shoot Density 4
Leaf:Shoot Ratio 5
Spatial Distribution 6

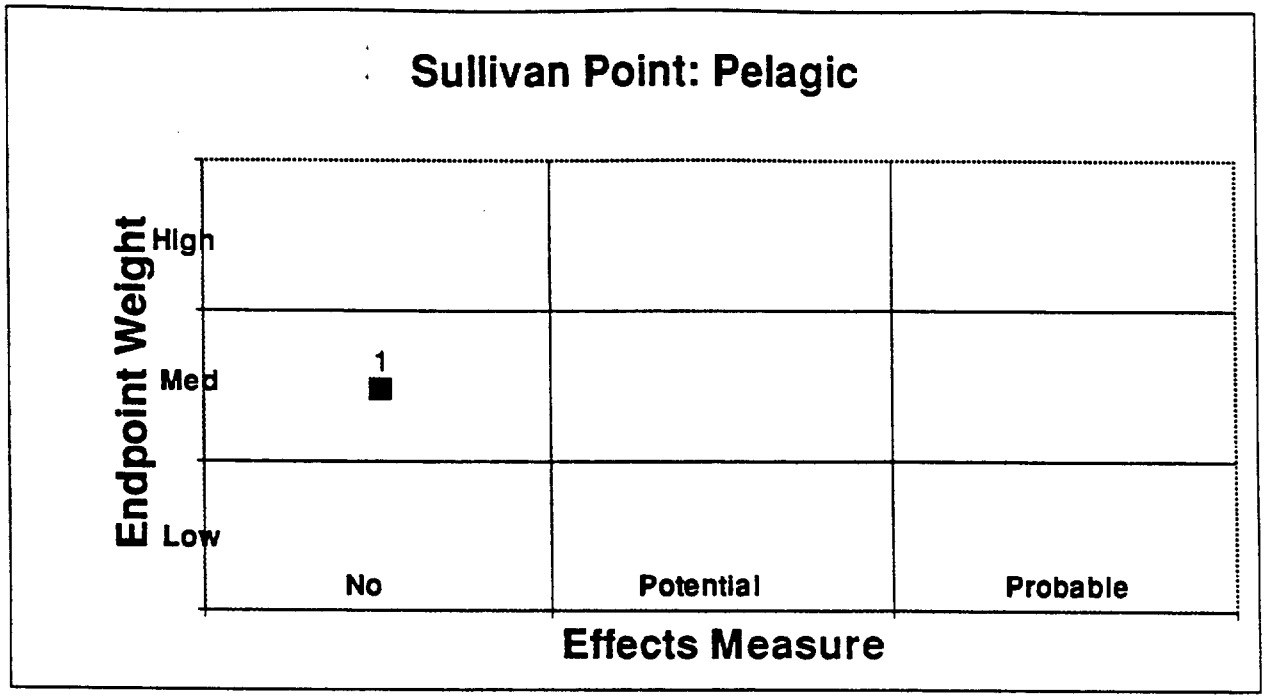
Potential/M
Due to lack of eelgrass at stations inside Clark Cove, a potential effect was assumed.



Sediment Bulk Chemistry A
Leaf Tissue Residue B
Root Tissue Residue C

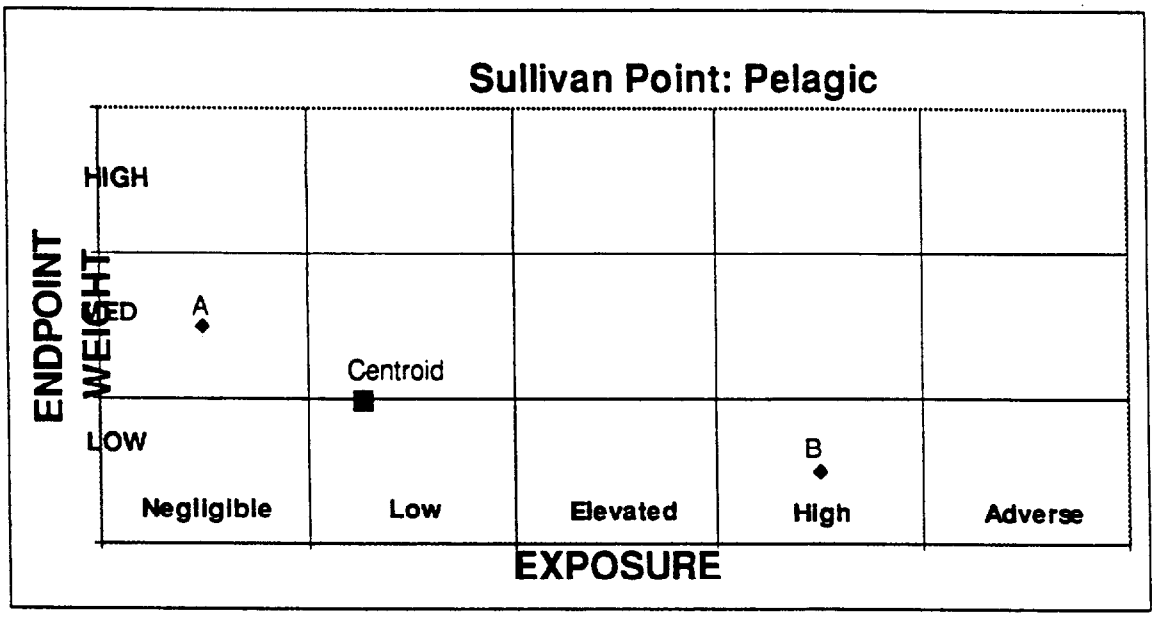
Elevated/M

Figure 7-4. Eelgrass assessment data for Clark Cove and weight of evidence from exposure and effects measures.



Arbacia Toxcity 1

No/M

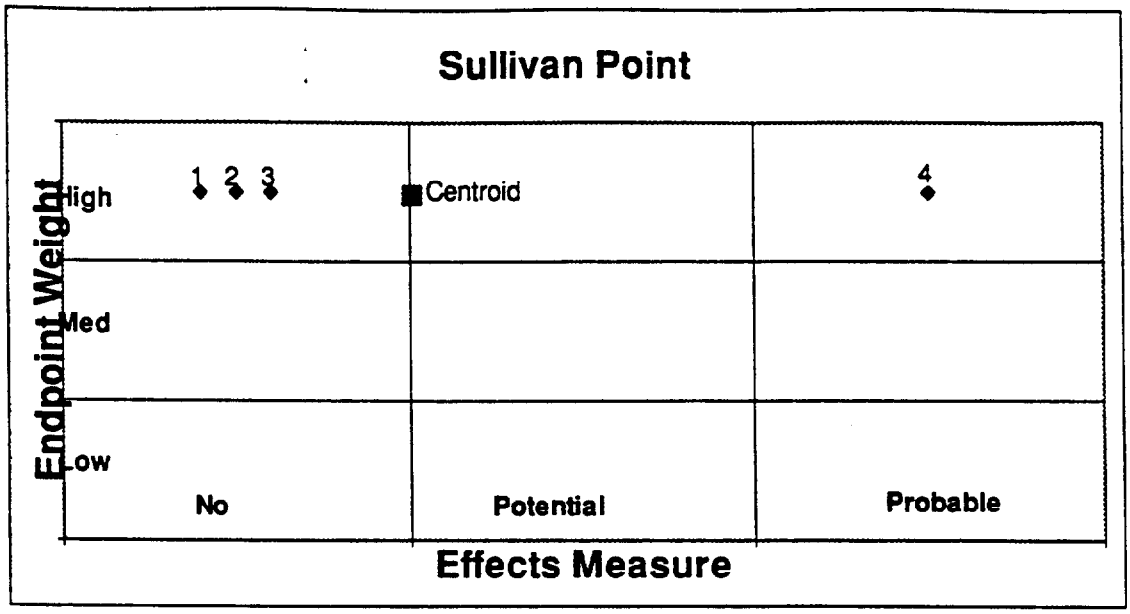


Surface Water A

Seep Water B

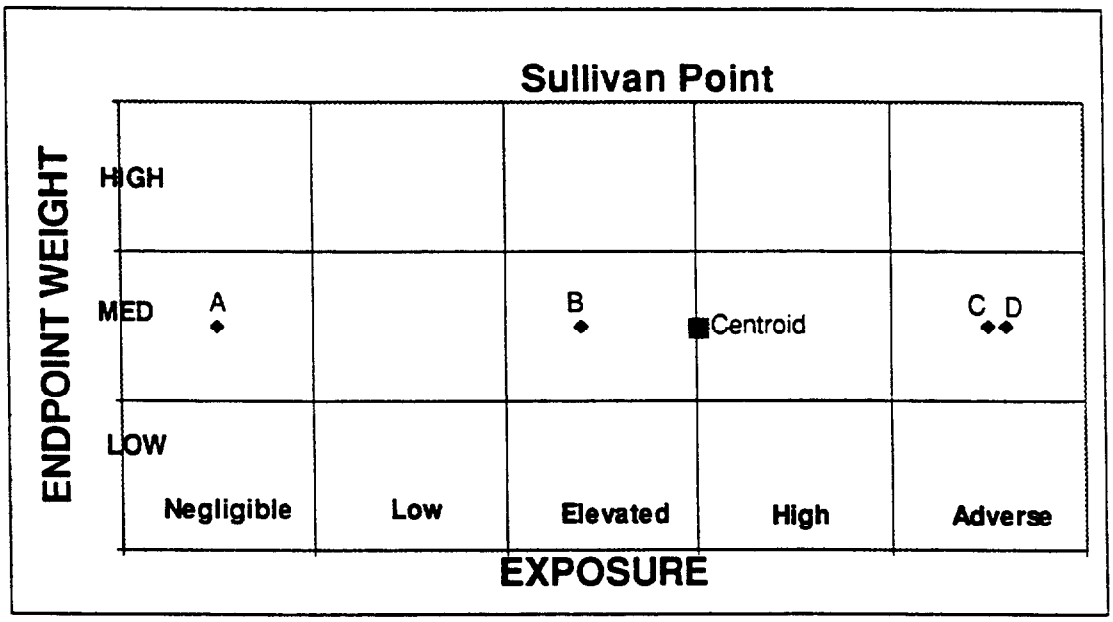
Low/L

Figure 7-6. Pelagic assessment data for Sullivan Point and weight of evidence from exposure and effects measures.



Species Evenness 1
Invertebrate Density 2
Species Richness 3
Amphipod Toxicity 4

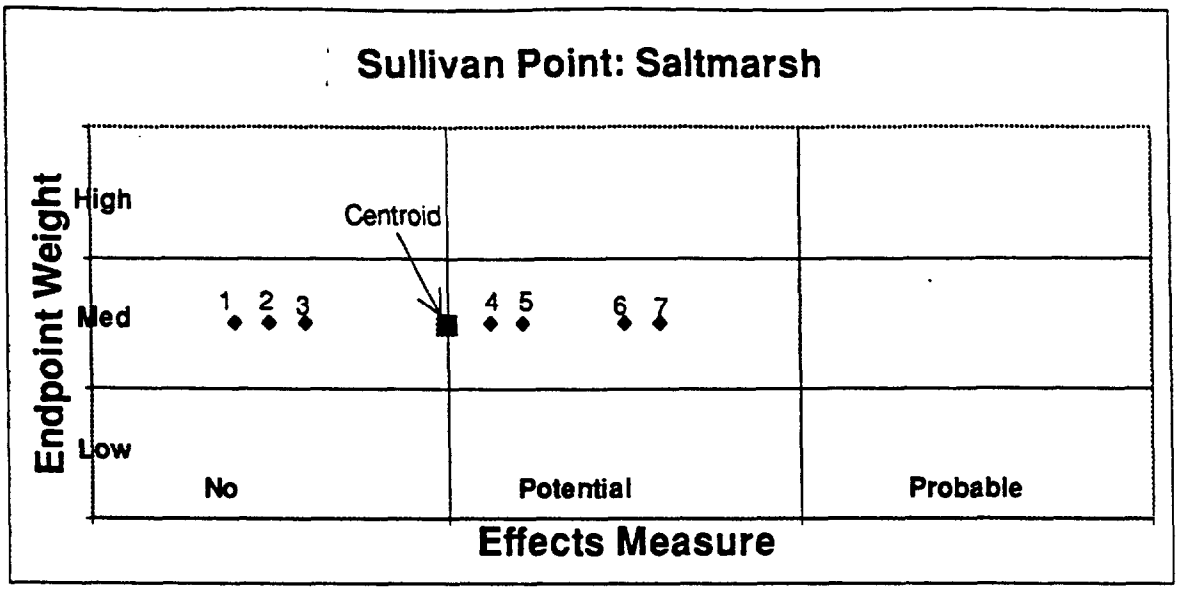
Potential/H



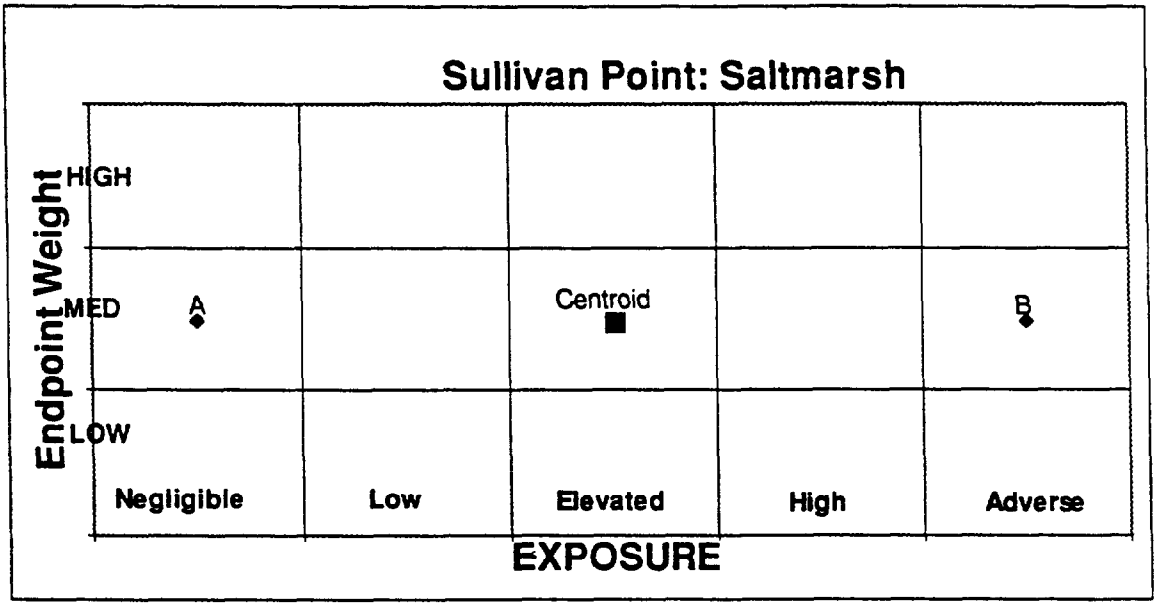
AVS-SEM A
Crustal Ratio B
Pore Water C
Bulk Sediment D

High/M

Figure 7-8. Benthic assessment data for Sullivan Point and weight of evidence from exposure and effects measures.

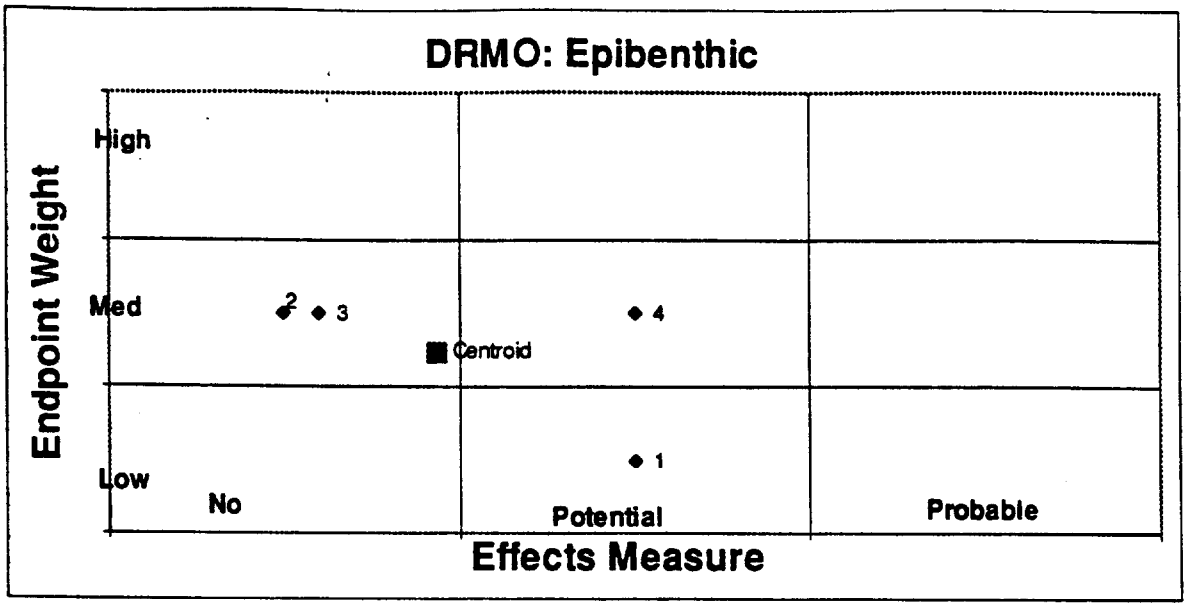


Spartina Cover 1
Littorina Live:Dead 2 Potential/M
Spartina Morphometerics 3
Number of Animal Taxa 4
Mollusc Abundance 5
Amphipod Abundance 6
Vascular Plant Cover 7

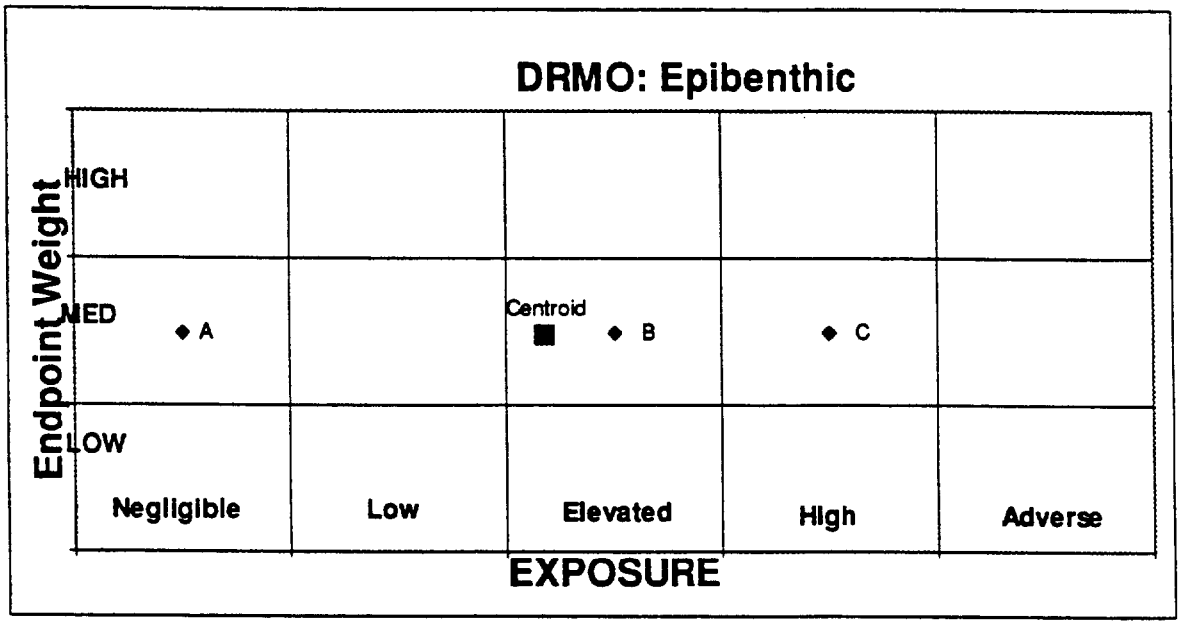


Spartina Tissue Residue A
Bulk Sediment Chemistry B Elevated/M

Figure 7-10. Salt marsh assessment data for Sullivan Point and weight of evidence from exposure and effects measures.

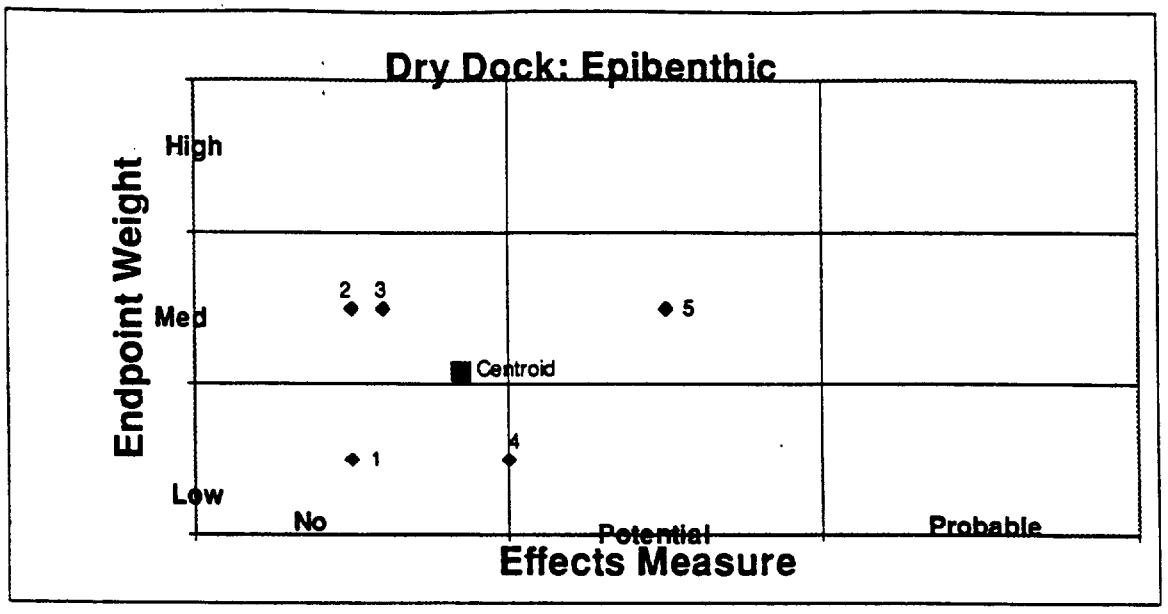


Fucoid Biomass 1
Mussel Condition Index 2 No/M
Mussel Density 3
Mussel Length 4

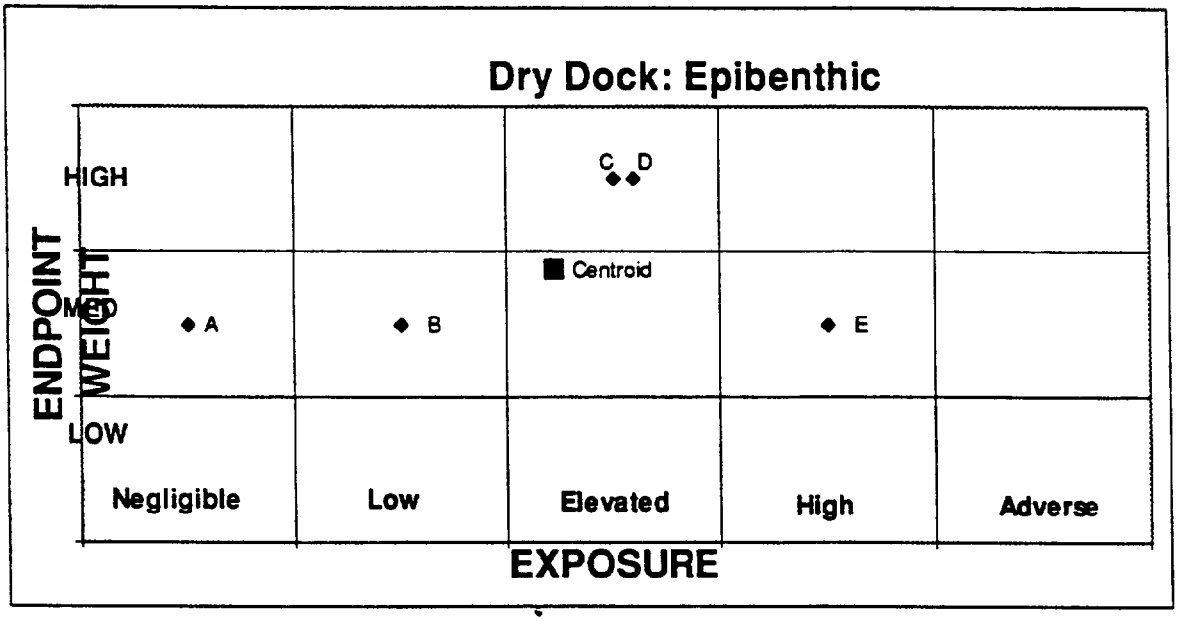


Surface Water A
Fucoid Algae Residue B Elevate/M
Mussel Residues C

Figure 7-12. Epibenthic assessment data for DRMO storage yard and weight of evidence from exposure and effects measures.

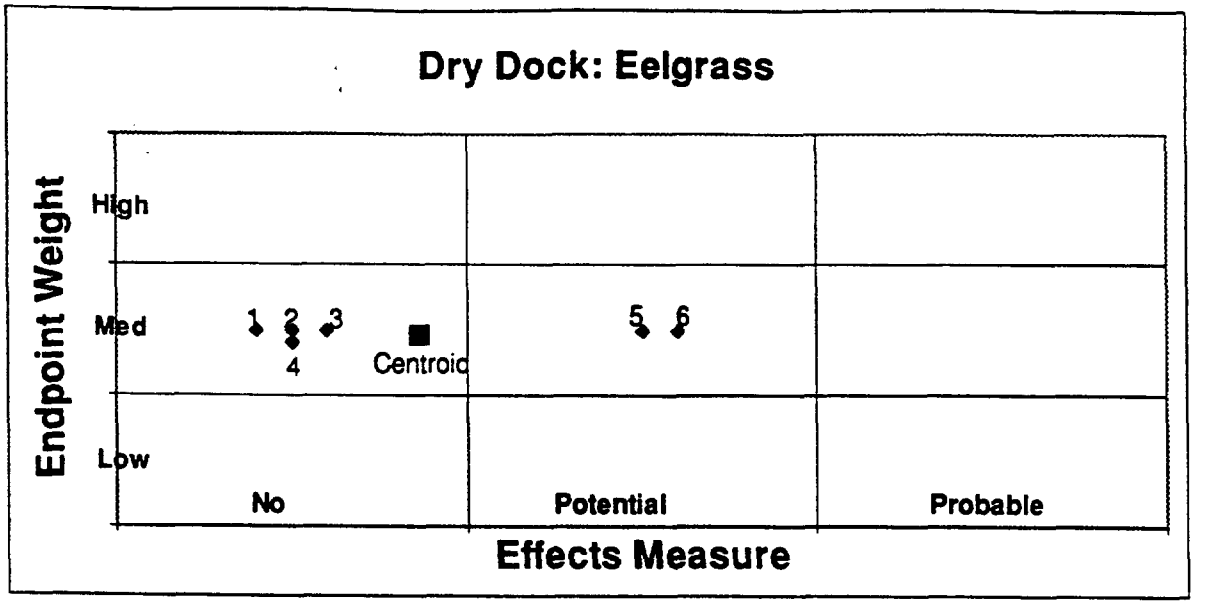


Lobster Density 1
Mussel Length 2
Mussel Condition Index 3 No/M
Fucoid Biomass 4
Mussel Density 5

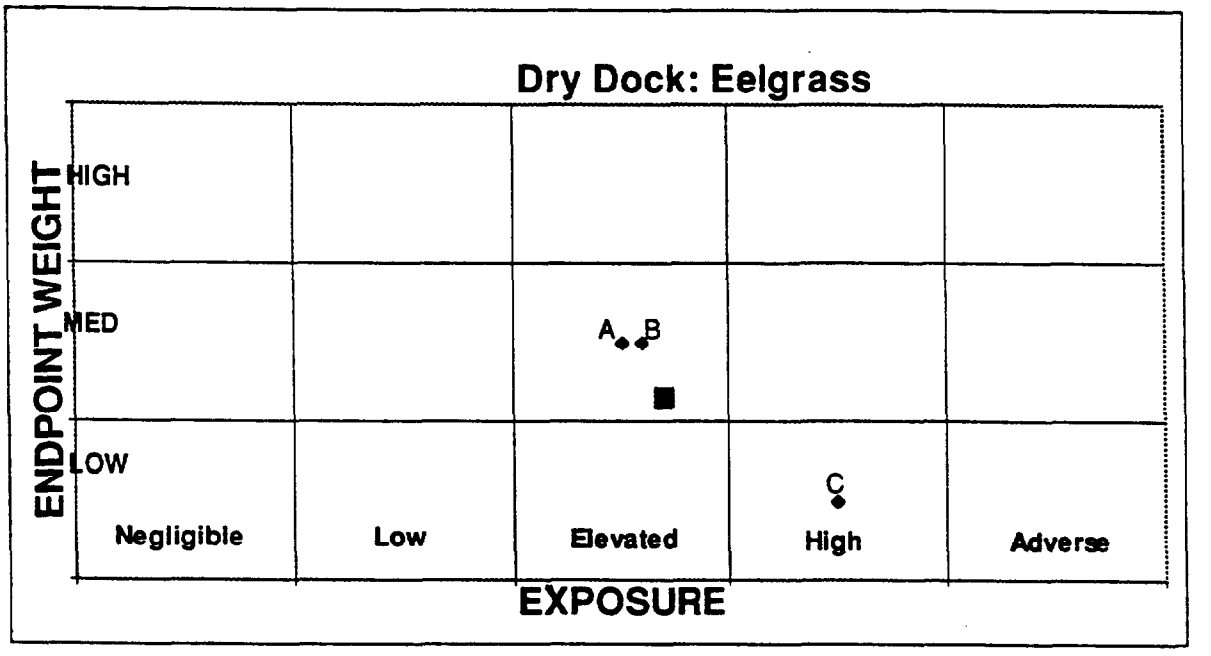


Surface Water A
Fucoid Residue B
Juvenile Lobster Hepatopancreas Residue C Elevate/M
Juvenile Lobster Tail+Claw Residues D
Mussel Residues E

Figure 7-14. Epibenthic assessment data for the Dry Dock AOC and weight of evidence from exposure and effects measures.

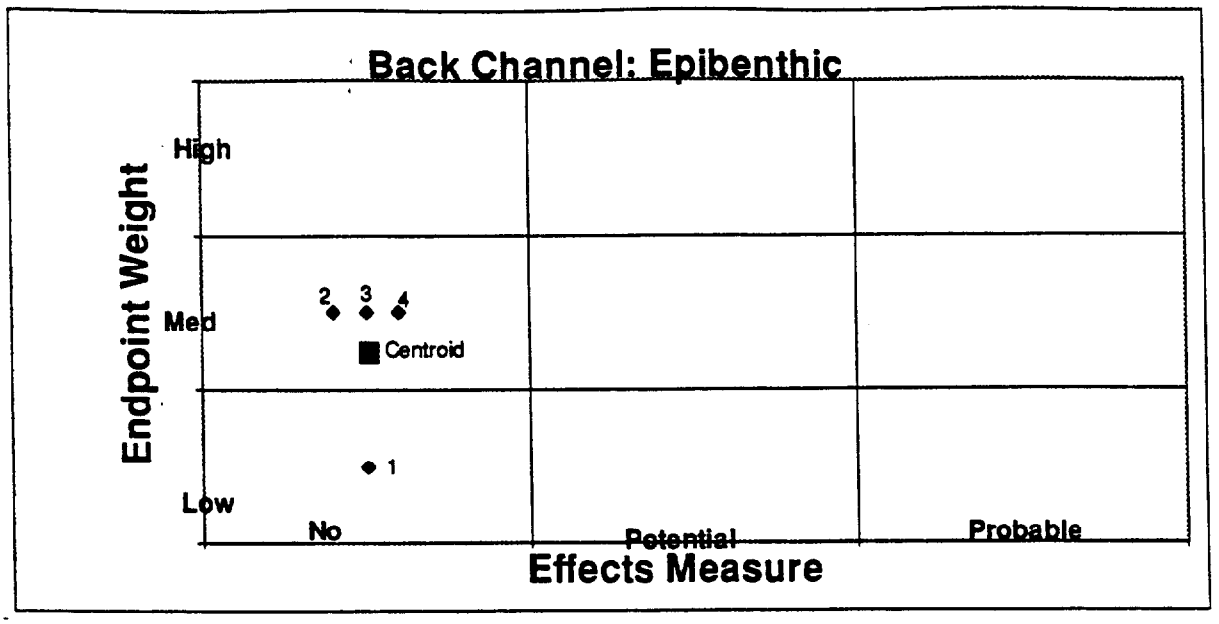


Spatial Distribution 1
Leaf:Shoot Ratio 2
Root Morphology 3 No/M
Leaf Morphology 4
Shoot Density 5
Reproductive Shoot Density 6

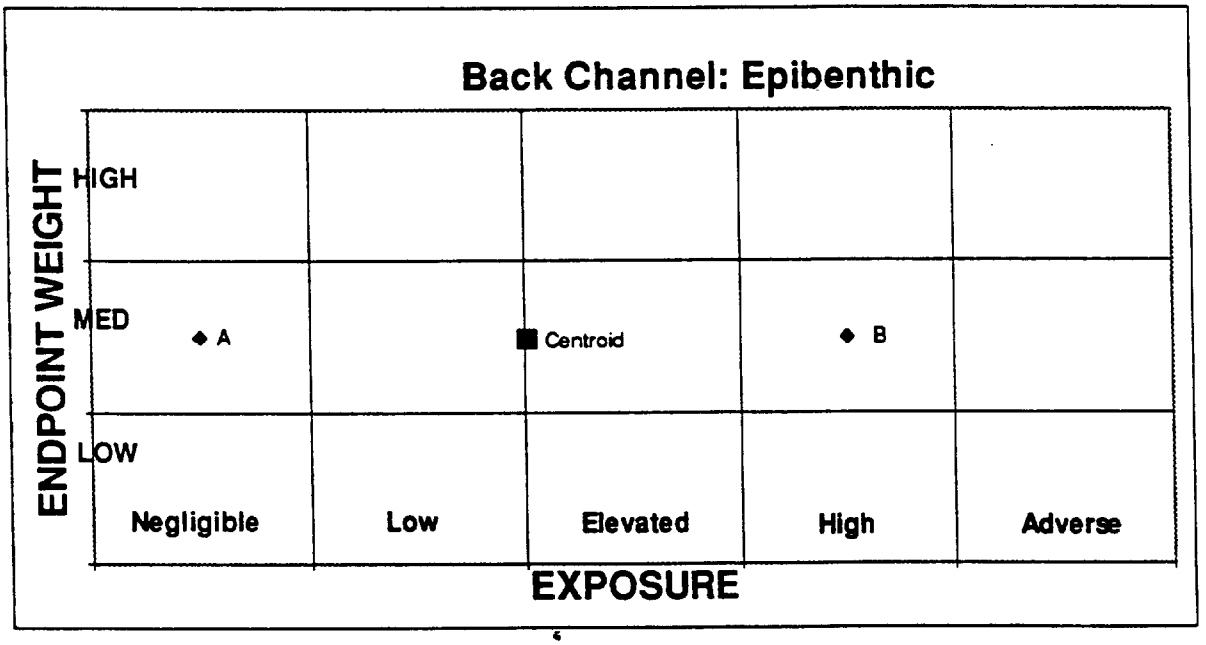


Eelgrass Leaf Tissue Residue A
Eelgrass Root Tissue Residue B Elevated/M
Bulk Chemistry C

Figure 7-16. Eelgrass assessment data for the Dry Dock AOC and weight of evidence from exposure and effects measures.

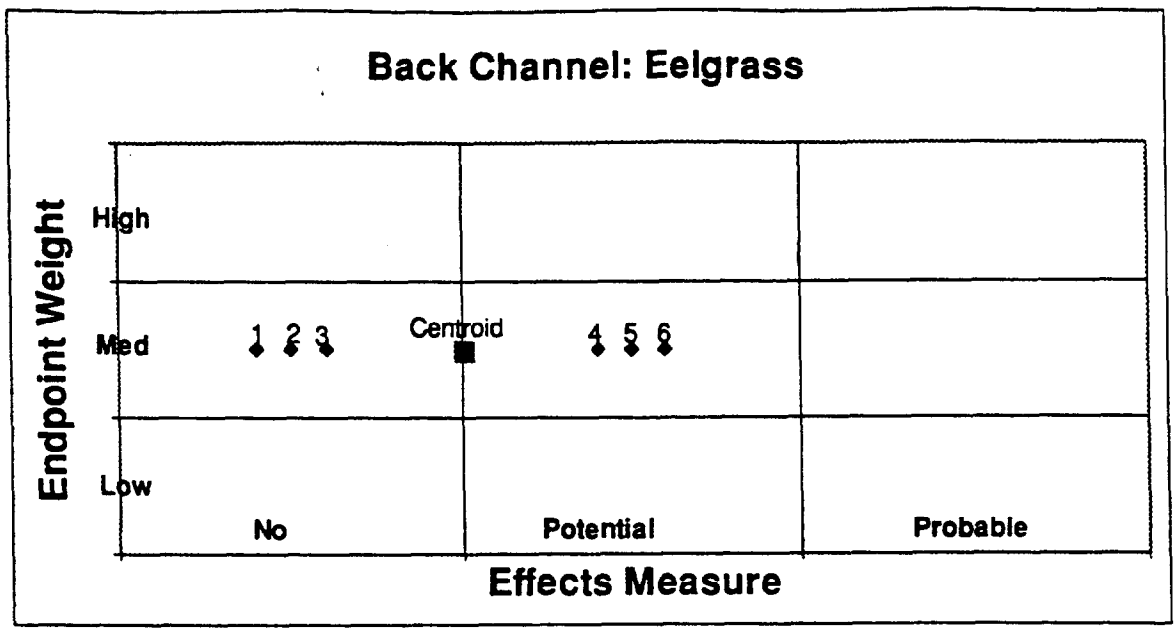


Lobster Abundance 1
Mussel Length 2 No/M
Mussel Density 3
Mussel Condition Index 4

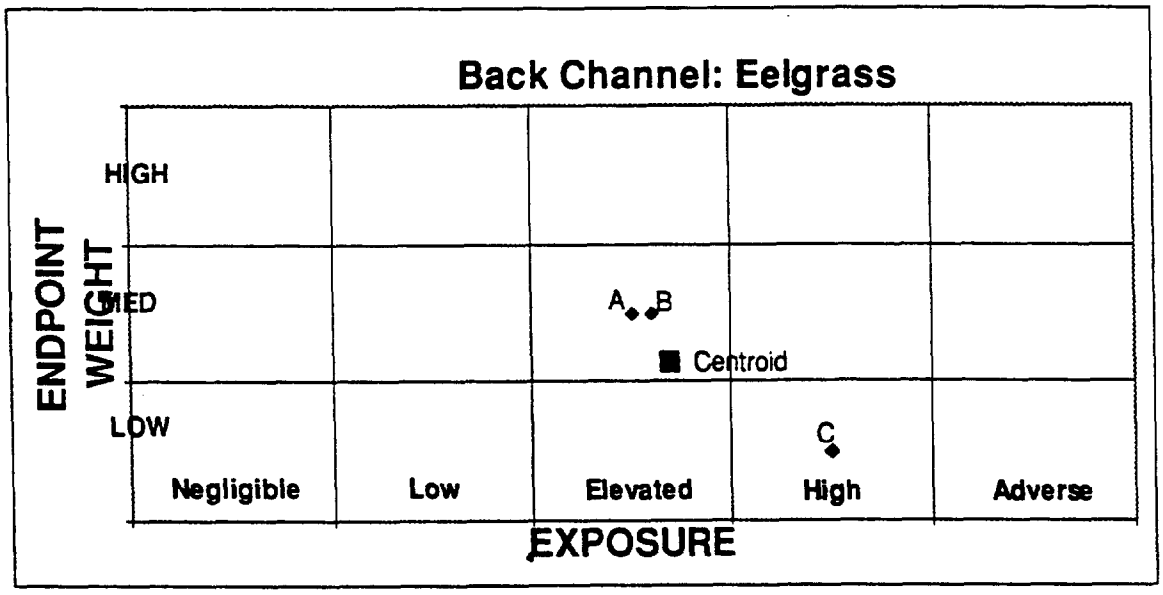


Surface Water A Elevated/M
Mussel Residues B

Figure 7-18. Epibenthic assessment data for the Back Channel and weight of evidence from exposure and effects measures.

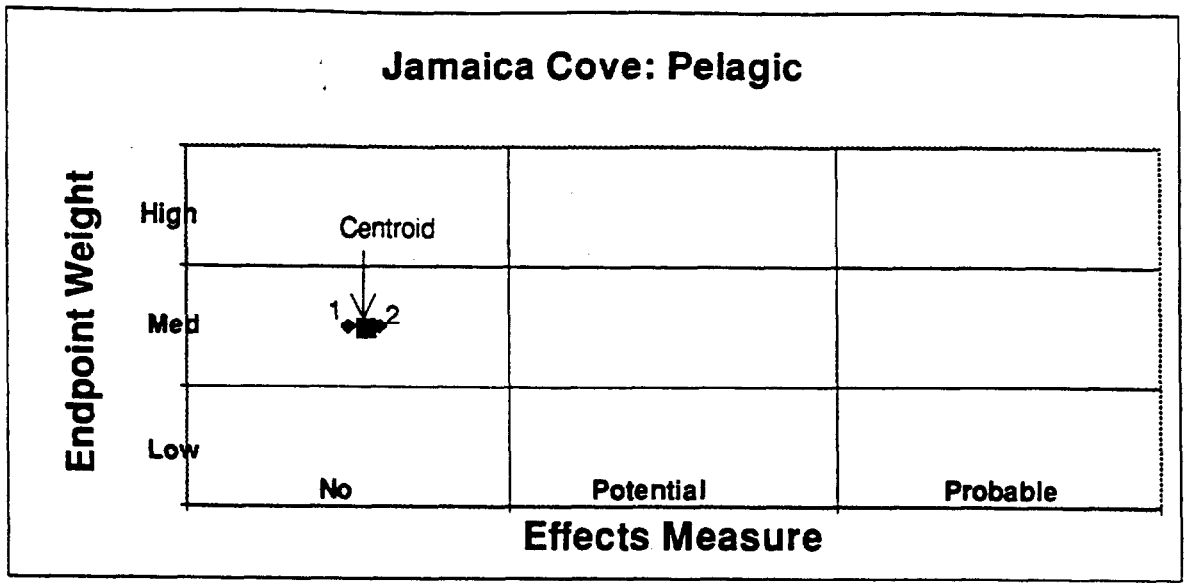


Shoot Density 1
Leaf:Shoot Ratio 2
Leaf Morphology 3 Potential/M
Root Morphology 4
Reproductive Shoot Density 5
Spatial Distribution 6

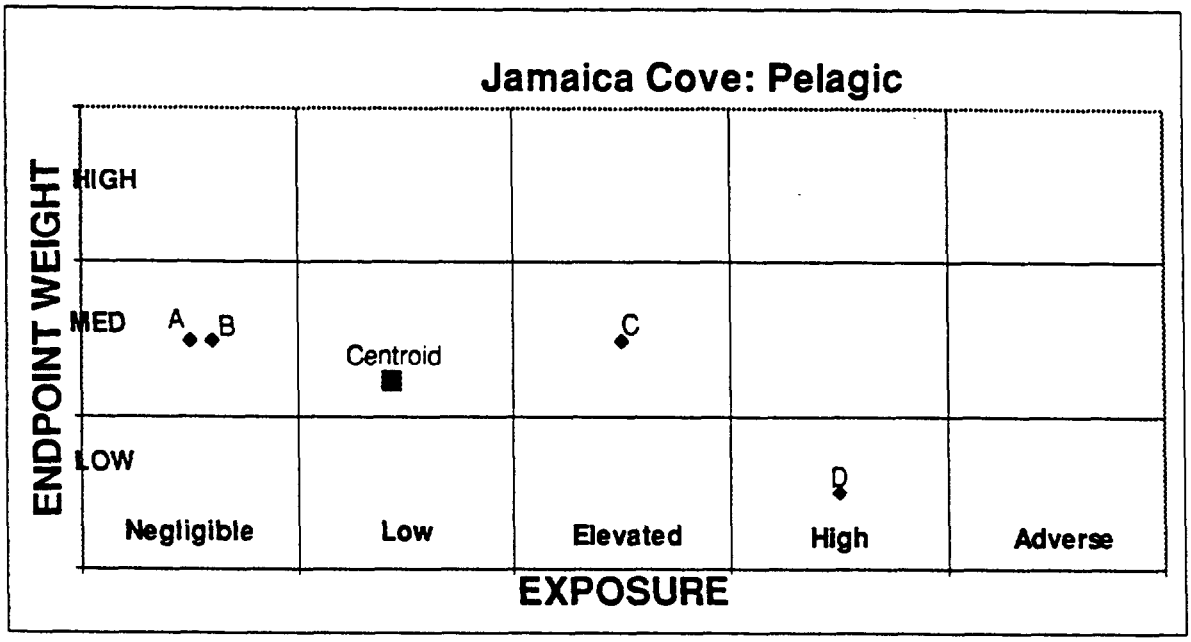


Leaf Tissue Residue A
Root Tissue Residues B
Bulk Chemistry C Elevated/M

Figure 7-20. Eelgrass assessment data for the Back Channel and weight of evidence from exposure and effects measures.

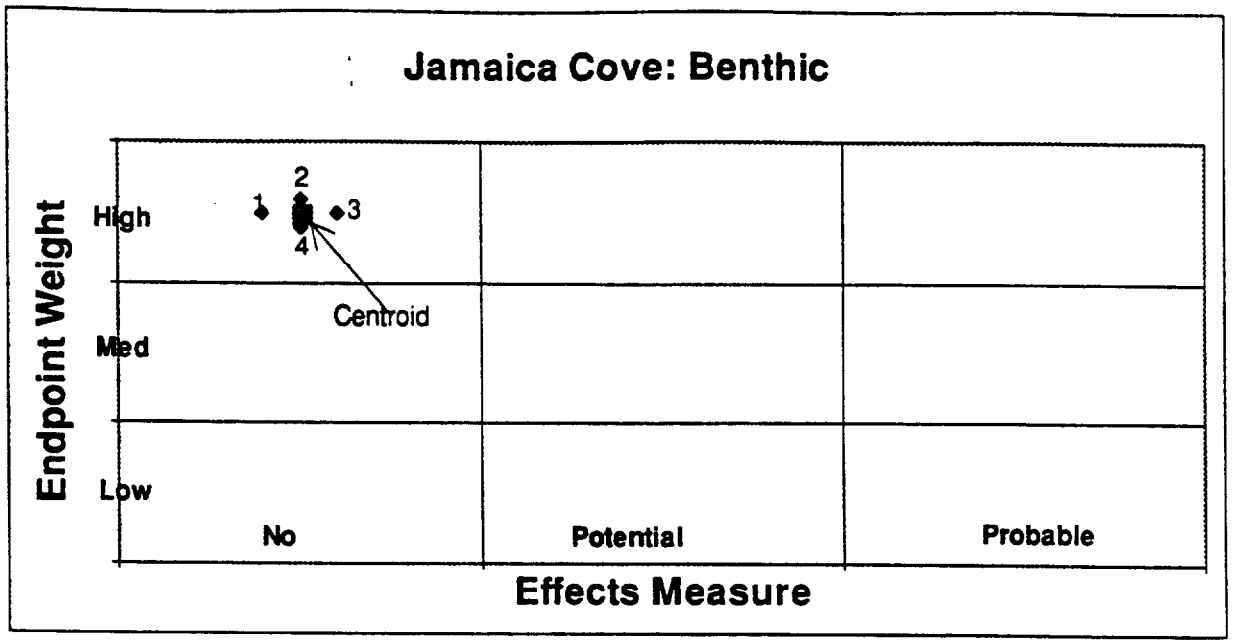


Deployed Mussel Scope for Growth (SFG) 1
Arbacia Toxcity 2 No/M

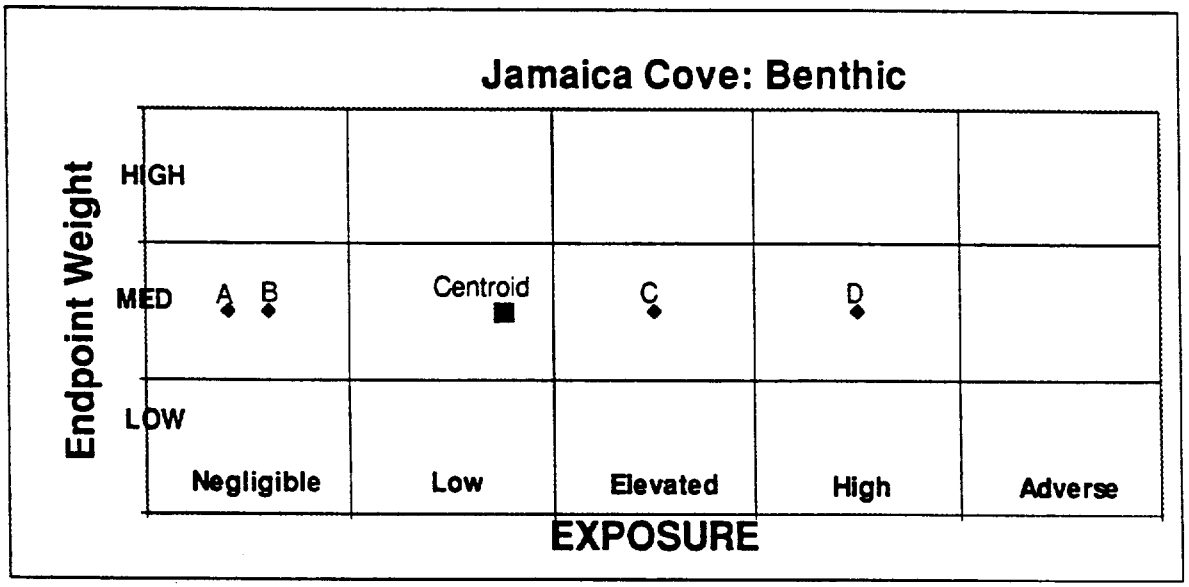


Surface Water A
Deployed Mussel Residues (PHASE I) B Low/M
Deployed Mussel Residue (PHASE II) C
Seep Water D

Figure 7-22. Pelagic assessment data for Jamaica Cove and weight of evidence from exposure and effects measures.

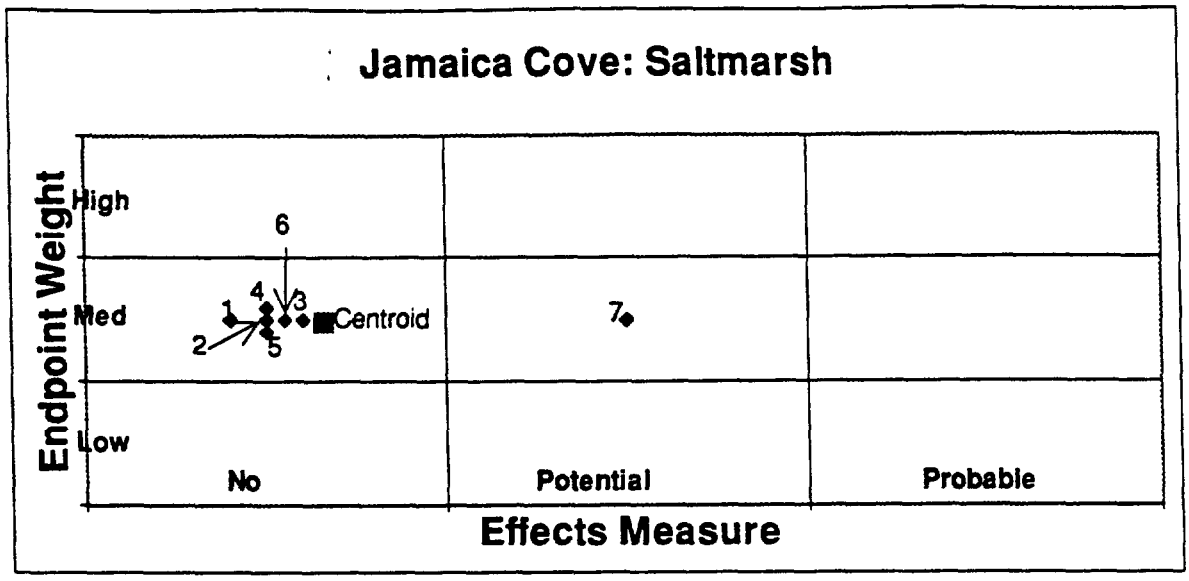


Invertebrate Density 1
Species Richness 2 — No/H
Species Evenness 3
Amphipod Toxicity 4

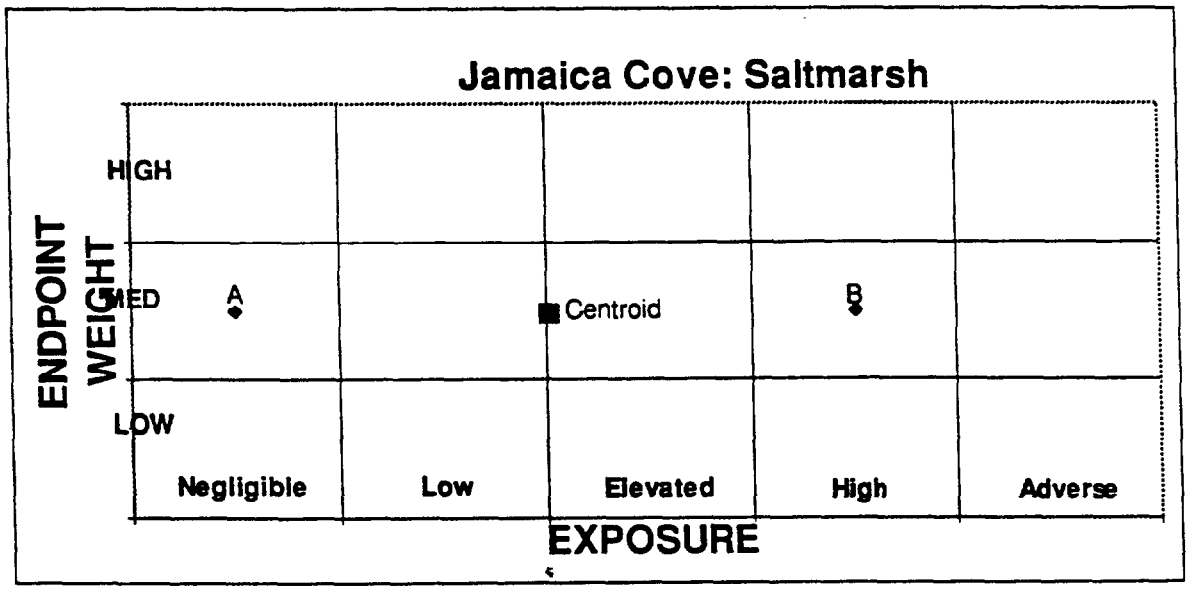


AVS-SEM A
Pore Water Toxic Unit B — Low/M
Crustal Ratio C
Bulk Sediment D

Figure 7-24. Benthic assessment data for Jamaica Cove and weight of evidence from exposure and effects measures.

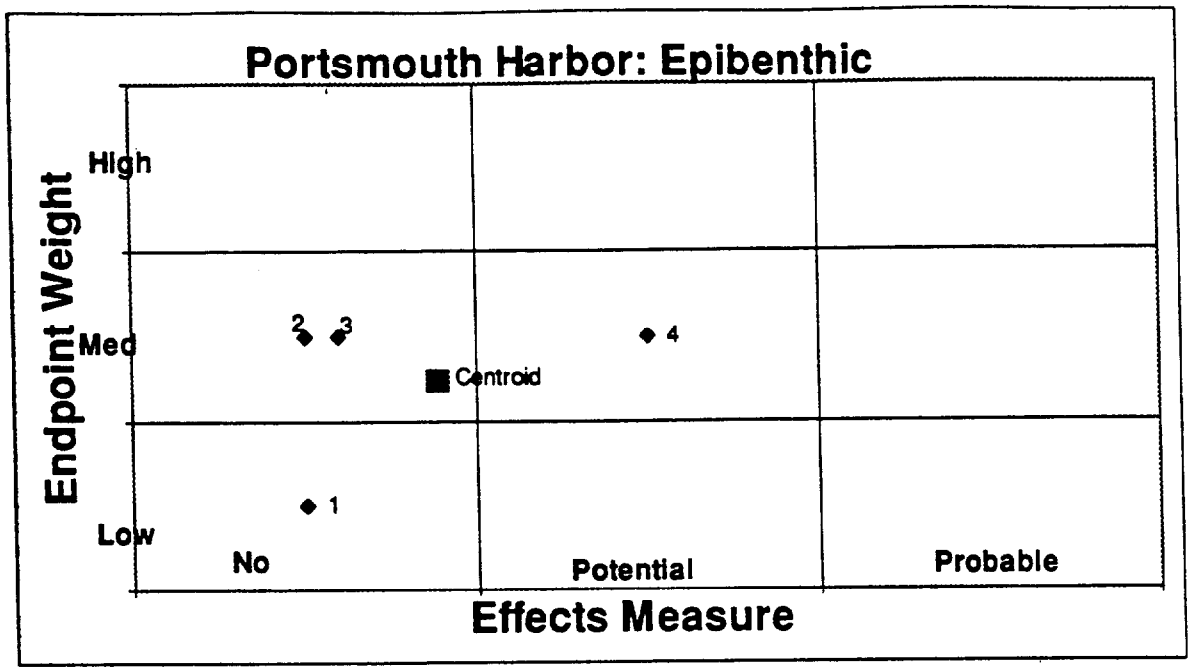


Spartina Cover 1
Spartina Morphology 2
Mollusc Abundance 3 No/M
Number of Animal Taxa 4
Vascular Plant Cover 5
Littorina Live:Dead 6
Amphipod Abundance 7

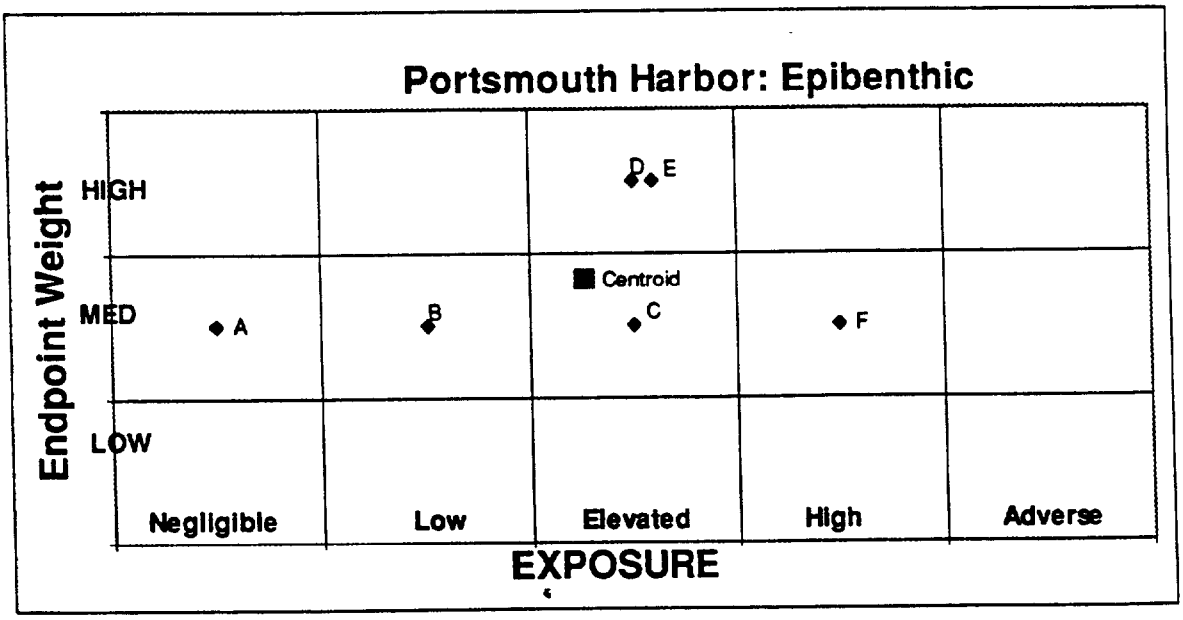


Spartina Residue A
Bulk Sediment B Elevated/M

Figure 7-26. Salt marsh assessment data for Jamaica Cove and weight of evidence from exposure and effects measures.



Loster Density 1
Mussel Density 2 No/M
Mussel Length 3
Mussel Condition 4



Surface Water A
Adult Lobster Tail Residues B
Adult Lobster Hepato C Elevated/M
Juvenile Lobster Hepato Res D
Juvenile Lobster Tail Res E
Mussel Residues F

Figure 7-28. Epibenthic assessment data for the Portsmouth Harbor focus area and weight of evidence from exposure and effects measures.

Appendix IX.3 Benthic Community

Patti Lynne Tyler
U.S. EPA Region I
Environmental Services Division
October 15, 1996
(updated May 4,.1998)

ASSESSMENT ENDPOINT: BENTHIC COMMUNITY AT PORTSMOUTH NAVAL SHIPYARD EXPLANATION FOR WEIGHT OF EVIDENCE TABLES

Introduction

As part of the risk characterization phase of the offshore ecological risk assessment for the Portsmouth Naval Shipyard, the assessment endpoint for the benthic community was evaluated against the following list of exposure and effect measures and the likelihood of an adverse impact on the benthic community has been determined by applying the weight of evidence approach (Menzie et al., 1996). The narrative for each of the measures support the weights identified in the attached tables.

MEASURES OF EFFECT	MEASURES OF EXPOSURE
Benthic Community Density (orgs/m2)	Comparison of sediment chemistry to crustal ration
Benthic Community Evenness (density/taxa)	Bulk Sediment Chemistry
Benthic Community Richness (# of taxa)	Sediment SEM/AVS chemistry
Ampelisca abdita sediment toxicity results	Pore water concentrations (EqP prediction)
	Sediment organic content*
	Sediment microbial indicator density*
	Sediment grain size*

* Sediment organic content and grain size are not directly evaluated as measures of exposure in the weight of evidence approach since this information was collected to determine depositional areas for sampling and they were not evaluated with respect in their relationship to the benthic community. The specific objective for the data collection efforts for the sediment microbial indicator density "was to determine potential existing sources and distribution of fecal contamination in Portsmouth Harbor waters" and the information can not be used to evaluate areas of concern related to disposal activities at the Portsmouth Naval Shipyard. Therefore these three measures were not evaluated in the weight of evidence approach in evaluating the potential risk to the benthic community.

The field sampling program for this effort consisted of sediment 23 stations targeted in depositional areas at the following locations:
- 21 stations in the lower Piscataqua (Figure 2-7 in Problem Formulation) 12 circumscribe Seavey Island with 6 of those 12 located in Clark Cove. Surface water and surface sediment samples for chemical analysis and toxicity testing were also collected at these same sampling locations. Please note that during the Phase II sampling efforts, 10a and 12a sampling locations were added but station 10a sampling efforts were limited to mussel and fucoid algae collection and 12a was limited to mussel and eelgrass sampling.

Grab samples and gravity cores were also collected at those same stations for size statistics, moisture, and organic content. Sediment texture distribution map was developed for all of the lower estuary, identifying the extent of the bottom habitat types.

1. Measure of Effect: Benthic Community Density, Richness and Evenness

Benthic Community Density (organisms/m2)
Benthic Community Richness (# of taxa)
Benthic Community Evenness (density/taxa)

Data Quality - High
It appears that all data quality objectives have been met.

Strength of Association between Assessment Endpoint and Measure of Effect - High
There is a direct biological relationship between these three measures and the assessment endpoint. Statistical analysis included the performance of multiple linear regressions for all 23 sites between total benthic species and mean sediment grain size, total PCBs , total metals, eelgrass biomass, and water toxicity. *"Stations with very low amphipod densities covered the entire range of sediment toxicities. However, only the sites with low toxicity had high amphipod densities, and no sites with high toxicity levels had high amphipod densities."* Total metals and PCBs only accounted for 17.6 % of the variation while sediment grain size and eelgrass biomass accounted for 55% of the variation. This measure is well accepted and is typically sensitive "assuming that the physical characteristics of the habitat (grain size, depth, current, etc.) remain reasonably constant, then changes in population parameters can often be associated with changes in chemical characteristics (nutrients, toxics, etc.)." These measures are based on direct observations and have been integrated with sediment toxicity and sediment chemistry

Study Design - Medium
Two sampling efforts were conducted, the first one in September 1991 and the second one in October 1992. High degree of variability with organismal density for stations and between replicates and sediment characteristics. Results are hard to interpret due to variable sediment characteristics since York Harbor sediments were very well to moderately sorted sand. In contrast, Clark Cove sediments were extremely poorly sorted mud. Because Station 10.5 (Storage Yard) was a non depositional area, sediments not sampled for this measure, and not included as a measure for this area of concern.

There was anomalous benthic community structure detected at 3 of the stations, 4 and 8 in Clark Cove and at Station 13 near the Dry Docks. However, stations 2, 4 and 8 had the greatest abundance with the fewest taxa and over 50% of the abundance dominated by the polychaete, *Streblospio benedicti*, which accounted for 76%, 58%, and 81% of the biomass, respectively. Station 13 had the lowest mean density of all stations.

There is disparity in infaunal populations within the Clark Cove 5 sampling locations when substrate and general location were similar and this provided the impetus for future sampling in this area as part of Phase II. Only five of the original 23 stations were sampled and there were large differences between Phase I and II density for stations 4 and 19. In addition, the sampling methodology differed from the use of four Shipek grab replicates = 0.008 m2. In 1992, sampling methodology consisted of five replicate cores = 0.04 m2.

Conclusions
The attached weight of evidence tables indicates a high endpoint weight for these three measures based on the weights applied to the attributes: data quality, strength of association between assessment endpoint and measure of effect and study design. All areas of concern associated with the Portsmouth Shipyard possessed densities of benthic organisms above the 5th percentile indicating no effect except for Station 13 located in the Dry Docks area in which the density were recorded below the 5th percentile resulting in a potential effect at this location. Station 13 had the lowest mean density of benthic organisms in comparison to all 23 stations sampled during the Phase I efforts.

All stations exceeded the 5th percentile for benthic community richness leading to an interpretation of no effect and potential effects to benthic community eveness were recorded below the 5th percentile at Station 13 in the Dry Docks area and above the 95th percentile range at Stations 4 and 8 located in Clark Cove.

"In sum, all analyses indicate that the benthic communities in Clark Cove were different from those in other areas. There was no clear indication that benthic communities at sites in the back channel or along the main channel near Seavey Island have been affected by contaminant releases from the Island. Reduction in total species at most of the Clark Cove sites particularly suggests environmental degradation in this area. However, variations in water and sediment contaminant measures only explained part of the reduction. Most was explained by variations in sediment grain size and eelgrass biomass." (Grizzle 1995)

2. Measure of Effect: *Ampelisca abdita* toxicity

This test was conducted as a 10 day solid phase test on a subsample of the sediment collected for the benthic community analysis. Testing was performed on sediments collected from the 23 Phase I sampling locations. The laboratory control sediment was collected from Long Island Sound-fine grained (>90% silt/clay), TOC ~ 2%. ERLN has an extensive database of no impacts associated with this reference sediment and during the testing, a concurrent water only reference test was conducted fulfilling quality assurance/quality control requirements.

Data Quality - High
All data quality objectives appear to have been met.

Strength of Association between Assessment Endpoint and Measure of Effect - High
There is a direct biological relationship between the results of the sediment toxicity test and the assessment endpoint since the test species, *Ampelisca abdita* is one of the dominant taxa identified during the benthic community analysis. Statistically significant reductions in survival were detected at 7 stations (9, 13, 16, 17, 18, 22 and 23). There was no positive correlation between any one single contaminant of concern and toxicity for those 7 stations. However, DDT exceeded screening levels at 5 of the 7 stations. Note that totals pesticides were not incorporated into the multiple linear regression models since it was not significantly correlated to variations in infaunal community abundance and diversity (Grizzle, 1995). Less than 20% survival was recorded from sediments collected from Station 9 (Sullivan Point) and Station 18 (Back Channel). Toxicity was associated with elevated metals and PAHs along with low field densities of benthic organisms.

Study Design - High
This test was only performed once during Phase I but benthic community analysis, analysis of sediment characteristics and chemical analysis were performed concurrently on the 23 Phase I sampling locations. The Storage Yard (10a) area of concern was not part of these Phase I stations and therefore this is not a measure for this area of concern.

Conclusions
The attached weight of evidence tables indicates a high weight for this endpoint based on the weights applied to the attributes: data quality, strength of association between assessment endpoint and measure of effect and study design. The interpretation of this measure is there is a probable toxic effect on the benthic community from sediments in Dry Docks, Sullivan Point, and Back Channel. There was significant reduction in survival at York Harbor reference locations when compared to the LIS sediment further questioning the use of York Harbor as an ideal reference location. Grain size may account for this toxicity since York Harbor grain size is ~ 0.5% and LIS~2.0%. However, the York Harbor stations had eelgrass, fucus, mussels, lobsters and flounder but amphipods were not abundant in the benthic samples. Low test survival was never associated with high field densities of benthic species and there was a correlation between TOC, grain size and survival and also elevated Pb and Cr at York River

3. Measure of Exposure: Comparison of Sediment Chemistry Data to Crustal Ratio

Data Quality - High
All data quality objectives appear to have been met.

Strength of Association between Assessment Endpoint and Measure of Exposure - Low
There is no biological linkage between crustal ratio and infaunal community and *"The degree of enrichment, or deviation from the expected metal concentration, indicates that there could be alternative sources of metal contributing to the observed distribution, presumably anthropogenic, but also possibly involving local geological mineral inputs or the atmospheric fall out of dust particles."* This leading to little correlation between chemical stressors originating from the shipyard and infaunal community response based on crustal ratios. Crustal ratio was integrated into a ranking scheme utilized to identify the most important contaminants of concern and benchmark does not appear to be very sensitive since it's application is only limited to metals. Metal enrichment could be a result of historic releases from the shipyard.

Study Design - Medium
Crustal ratio measures limited to Main Channel, Back Channel and Clark Cove. It is unclear as to what stations comprise Main Channel and Back Channel. Storage Yard, Drydocks, Sullivan Point and Jamaica Island are areas of concern that are not evaluated with respect to this measure. This lack of data for those areas of concern make it hard to make comparisons with other measures. It is also difficult to determine the specifics of the standard method employed and the sensitivity of the endpoint since core depth varied with location from surface to 130 cm deep and data is not presented in the document.

Conclusions
The attached weight of evidence tables indicates a medium weight for this endpoint based on the weights applied to the attributes: data quality, strength of association between assessment endpoint and measure of effect and study design. Crustal ratios from Storage Yard, Clark Cove and Back Channel indicate

enrichment of metals above the predicted ratio leading to an interpretation of low exposure to the assessment endpoint from this measure.

4. Measure of Exposure: Bulk Sediment Chemistry
a. Comparison to NOAA ER-L Values

Data Quality - High
All data quality objective appear to have been met.

Strength of Association between Assessment Endpoint and Measure of Exposure- Medium
Based on the NOAA database, the NOAA ER-L concentration for a specific contaminant represents the overlap of various biological effects data (lab tests, field tests and modeling) with chemical concentrations. The ER-L represents the 10th percentile of data that demonstrates a biological response. The range of contaminant concentrations between the ER-L and ER-M represent possible effects, *"The range in concentrations over which toxic effects are occasionally observed (Long et al. 1995)."* The response is not well correlated with duration of exposure and this measure does not take into account bioavailability or bioaccumulation. Guidelines are applicable to interpretation of environmental data since they are generated from field and lab studies and theoretical modeling.

Study Design - Medium
Chemical analysis, benthic community analysis and sediment toxicity testing were conducted on the same Phase I sediment samples . Sediment sampling and chemical analysis were conducted once. This measure is not site specific and does not take into factors that may influence bioavailability.

Conclusions
The attached weight of evidence tables indicates a medium weight for this endpoint based on the weights applied to the attributes: data quality, strength of association between assessment endpoint and measure of effect and study design.

The results of the chemical analysis and comparison to NOAA ER-L values indicate the following exceedances which leads to the interpretation of low exposure for many compounds in sediments at most of the areas of concern:

*Chromium concentrations exceeded ER-L value in Clark Cove
*Copper exceeded ER-L at Dry Dock (station 12)
*"ER-L toxicity values for lead were exceeded at the following stations: Clark Cove (stations 4,6,7, and 8); near the Dry Docks (stations 10, 12 and 17) and ; Back Channel (stations 18 and 19)
*Exceedance of nickel ER-L values were detected at Clark Cove (stations 5,6,7, and 8) and Dry Dock (station 10)
*Mercury ER-L values exceeded at Clark Cove (stations 4,5,6,7, and 8); dry docks (stations 10, 12 and 17); back channel (station 19);
*Silver ER-L exceeded in Clark Cove (station 7)
*Zinc ER-L exceeded at Clark Cove (stations 7 and 8);
*ER-L for total PCBs was exceeded at Clark Cove (station 7 and 8); dry dock (station 12 and 10); back channel (station 18)
*ER-L values for total DDT was exceeded at most all stations

b. Comparison of Sediment Chemistry to NOAA ER-M Values

Data Quality - High
Same as above

Strength of Association between Assessment Endpoint and Measure of Exposure - Medium
Same as above but the ER-M value represents the median percentile in which biological response has occurred. *"The toxicity threshold values were used to identify station locations where sediment contamination levels could be toxic to marine organisms."* Contaminant concentrations in sediments that exceed the ER-M may be frequently responsible for toxic effects. *"The concentrations equivalent to and above the ER-M value represent a probable-effects range within which effects would frequently occur"* (Long et al., 1995).

Study Design - Medium
Same as above

Conclusions
The attached weight of evidence tables indicates a medium weight for this endpoint based on the weights applied to the attributes: data quality, strength of association between assessment endpoint and measure of effect and study design. There is an adverse exposure of sediment contaminants to benthic organisms in sediments collected from Sullivan Point, Dry Docks, Clark Cove, and Back Channel.

Results of sediment chemical analysis in comparison to NOAA ER-M values include:
Clark Cove: exceeded ERM for Hg, tPCB, ACHLOR, DDT, and DDD.
Sullivan Point: exceeded ERM for DDT and Ni.
Dry Docks: exceeded ERM for Cu, Hg, Ni, Pb, Zn, FLRAN, pyrene, BAA, chrysene, BAP, DIBAHA, sumPAH, HPAHs, tPCB, ACHLOR, DDD, and DDT.
Back Channel: -exceeded ERM for PHEN, fluorene, ANTH, pyrene, BAA, BAP, HPAHs, and DDD.
Jamaica Cove: only exeedance of ERM was two deep core samples (20-28 cm and 50-58 cm) for DDT.

c. Comparison of Sediment Chemistry to State of Washington Cleanup Levels

Data Quality - High
All data quality objectives appear to have been met.

Strength of Association between Assessment Endpoint and Measure of Exposure - Low
There is no biological relationship between the WA-CL approach and the infaunal community at Portsmouth Naval Shipyard areas of concern since WA-CL values are not specific to this geographic region.
Study Design - Medium

Conclusions
The attached weight of evidence tables indicates a low weight for this endpoint based on the weights applied to the attributes: data quality, strength of association between assessment endpoint and measure of exposure and study design. Concentrations of total DDT exceeded WA-CL at Clark Cove (station 8), Sullivan Point (station 9), Dry Dock (station 12) and Back Channel (station 18).

Note: The consensus from the WOE analysis was that bulk sediment chemistry should be evaluated as one measure of exposure and the magnitude of contamination should be evaluated as follows:

≤ reference condition or below conservative benchmark (ER-L) = negligible exposure
> qualitative screening level = low exposure
statistically > reference condition = elevated exposure
> a conservative benchmark (ER-L) = high exposure
> a nonconservative benchmark (ER-M) = adverse exposure

5. Measure of Exposure: Sediment SEM/AVS Chemistry

Data Quality - High
All data quality objectives appear to have been met but SEM/AVS only analyzed in Clark Cove, Dry Docks, Jamaica Cove and Sullivan Point sediments

Strength of Association between Assessment Endpoint and Measure of Exposure - High
This measure is a good predictor of whether divalent metals (Cu, Cd, Ni, Zn and Pb) are bioavailable to the infaunal species.

Study Design - Low
SEM/AVS was only measured in Clark Cove, Dry Docks, Jamaica Cove and Sullivan Point sediments and lead amended sediments during Phase II (summer) . There was no evaluation of seasonal variability since this analysis was only completed in the late summer when AVS would be expected to be high due to increased primary productivity. Worst case scenario would include AVS measurements in winter.

Conclusions
The attached weight of evidence tables indicates a low weight for this endpoint based on the weights applied to the attributes: data quality, strength of association between assessment endpoint and measure of exposure and study design. All SEM/AVS ratios were less than 1 indicating a low exposure of divalent metals to benthic organisms.

6. Pore water concentration based on equilibrium partitioning (EqP)

Data Quality - High
All data quality objectives appear to have been met.

Strength of Association between Assessment Endpoint and Measure of Exposure - Medium
This measure is a good predictor of whether organic compounds are bioavailable to the infauna species.

Study Design - Medium
Chemical analysis of organic contamiants, total organic carbon, benthic community analysis and sediment toxicity testing were conducted on the same Phase I sediment samples. Sediment sampling and chemical analysis were conducted once. By using the EqP approach this measure take into account site-specific factors that may influence bioavailability. However, no direct measurement of pore water concentrations were performed.

Conclusions

The attached weight of evidence tables indicates a medium weight for this endpoint based on the weights applied to the attributes: data quality, strength of association between assessment endpoint and measure of exposure, and study design. Predicted pore water concentrations were normalized to toxic units (TUs) by dividing the predicted pore water concentration by the ambient chronic water quality criteria or LC50 value for the most sensitive species. If the TU calculated for any of the surface grab sediment samples was greater than one for any COC, adverse exposure was assumed. Toxic units greater than one were calculated for PHEN, FLUOR, ANTH, FLURAN, Pyrene, and DDT in the Back Channel,: PHEN and FLUORAN in the Dry Docks; and DDT at Sullivan Point.

PORTSMOUTH NAVAL SHIPYARD
ASSESSMENT ENDPOINT INFORMATION SUMMARY

ASSESSMENT ENDPOINT: Benthic- Jamaica Cove

MEASUREMENT ENDPOINT	DATA QUALITY	STRENGTH OF RELATIONSHIP	STUDY DESIGN	ENDPOINT WEIGHT	OUTCOME	INTERPRETATION	ISSUES
Measure of Effect: Benthic Community Density	High	High	Medium	High	similar to reference, above 5th percentile	no effect	
Measure of Effect: Benthic Community Richness	High	High	Medium	High	similar to reference, above 5th percentile	no effect	
Measure of Effect: Benthic Community Evenness	High	High	Medium	High	similar to reference, above 5th percentile	no effect	
Measure of Effect: A. abdita toxicity	High	High	High	High	No toxicity relative to the control sediment	no effect	- Grain size control, reference site
Measure of Exposure: Bulk Sediment Chemistry	High	Medium	Medium	Medium	> ERL	high exposure	-Station 19 exceeded the following NOAA ERLs: pyrene, phenanthrene, benzo-a-anthracene, anthracene, dibenz(a,h)anthracene, total PCBs., As, Cr, Pb, Hg, Ni; -only exeedance of ERM was two deep core samples for DDT.
Measure of Exposure Crustal Ration	High	Low	Medium	Medium	metal concentrations enriched	elevated exposure	Ag, Hg, Pb, Cr, and Cd were enriched above crustal background
Measure of Exposure: SEM/AVS	High	High	Low	Medium	AVS >> SEM	negligilbe exposure	Seasonal variation in AVS and SEM was not measured
Measure of Exposure: EqP Pore water Toxic Unit (TU)	High	Medium	Medium	Medium	< TU	negligible exposure	

PORTSMOUTH NAVAL SHIPYARD
ASSESSMENT ENDPOINT INFORMATION SUMMARY

ASSESSMENT ENDPOINT: Benthic - Sullivan Point

MEASUREMENT ENDPOINT	DATA QUALITY	STRENGTH OF RELATIONSHIP	STUDY DESIGN	ENDPOINT WEIGHT	OUTCOME	INTERPRETATION	ISSUES
Measure of Effect: Benthic Community Density	High	High	Medium	High	similar to reference, above 5th percentile	no effect	
Measure of Effect: Benthic Community Richness	High	High	Medium	High	similar toreference, above 5th percentile	no effect	
Measure of Effect: Benthic Community Evenness	High	High	Medium	High	similar to reference, above 5th percentile	no effect	
Measure of Effect: A. abdita toxicity	High	High	High	High	statistically higher toxicity relative to control	probable effect	Toxicity associated with high DDT Survival <20% and below minimum detectable difference - Grain size control reference site.
Measure of Exposure: Bulk Sediment Chemistry	High	Medium	Medium	Medium	above ERM for DDT and Ni above ERL for PAHs and metals	adverse exposure	-Station 9 exceeded the following NOAA ER-Ls: pyrene, phenanthrene, Baanthracene, anthracene, flourene, total pest., DDT, Pb, Hg, Zn, Ni, Cu - Sta. 9 exceeded ERMs for DDT and Ni.
Measure of Exposure AVS-SEM	High	High	Low	Medium	AVS >> SEM	negligible exposure	no seasonal data for AVS and SEM
Measure of Exposure Crustal Ratio	High	Medium	Medium	Medium	< ERM	low exposure	
Measure of Exposure: EqP pore water toxic unit (TU)	High	Medium	Medium	Medium	> TU for DDT	adverse exposure	DDT exceeded pore water TU

PORTSMOUTH NAVAL SHIPYARD
ASSESSMENT ENDPOINT INFORMATION SUMMARY

ASSESSMENT ENDPOINT: Benthic - Dry docks

MEASUREMENT ENDPOINT	DATA QUALITY	STRENGTH OF RELATIONSHIP	STUDY DESIGN	ENDPOINT WEIGHT	OUTCOME	INTERPRETATION	ISSUES
Measure of Effect: Benthic Community Density	High	High	Medium	High	less than 5th percentile	potential effect	station 13 had the lowest mean density of all 23 stations
Measure of Effect: Benthic Community Richness	High	High	Medium	High	similar to reference, above 5th percentile	No effect	highest # of taxa at station 12
Measure of Effect: Benthic Community Evenness	High	High	Medium	High	similar to reference, above 5th percentile	no effect	
Measure of Effect: A. abdita toxicity	High	High	High	High	statistically higher toxicity relative to control	probable effect	40% survival at station 13
Measure of Exposure: Bulk Sediment Chemistry	High	Medium	Medium	Medium	> ERL for metals and organics > ERM for metals and organics	adverse exposure	-Exceeded ERLs for: flouranthene, phenanthrene, pyrene, chrysene, BAP, Benzo(a)anthracene, anthracene, flourene, dibeno(a,h)anthracene, total PCBs, total pest., DDT, Pb, Hg, ZN, Ni, Cu. -Exceeded ERMs for Cu, Hg, Ni, Pb, Zn, FLRAN, Pyrene, BAA, Chrysene, BAP, DIBAHA, sumPAH, HPAH, tPCB, ACHLOR, DDD, and DDT.
Measure of Exposure; Crustal Ratio	High	Low	Medium	Medium	enriched	elevated exposure	As, Ag, Cd, Cr, Cu, Pb, and Zn enriched above crustal ratio
Measure of Exposure: EqP Pore Water Toxic Unit	High	Medium	Medium	Medium	> TU	adverse exposure	PHEN and FLURAN exceeded TU
Measure of Exposure: SEM/AVS	High	High	Low	Medium	AVS >> SEM	negligible exposure	no seasonal data on AVS and SEM

PORTSMOUTH NAVAL SHIPYARD
ASSESSMENT ENDPOINT INFORMATION SUMMARY

ASSESSMENT ENDPOINT: Benthic - Back Channel

MEASUREMENT ENDPOINT	DATA QUALITY	STRENGTH OF RELATIONSHIP	STUDY DESIGN	ENDPOINT WEIGHT	OUTCOME	INTERPRETATION	ISSUES
Measure of Effect: Benthic Community Density	High	High	Medium	High	similar to reference, above 5th percentile	No effect	
Measure of Effect: Benthic Community Richness	High	High	Medium	High	similar to reference, above 5th percentile	No effect	
Measure of Effect: Benthic Community Evenness	High	High	Medium	High	similar to reference, above 5th percentile	No effect	
Measure of Effect: A. abdita toxicity	High	High	High	High	statistically higher toxicity relative to control	Probable Effect	Elevated level of Pb, sum PAH, Cd and Cu associated with tox. Survival <20% at station 18
Measure of Exposure: Crustal ratio	High	Low	Medium	Medium	enriched	elevated exposure	Pb and Hg enriched at least > 3 times upper bound and CD > than 10 times upper bound
Measure of Exposure: Bulk Sediment Chemistry	High	Medium	Medium	Medium	> ERL for PAHs > ERM for PAHs	adverse exposure	-Exceeded NOAA ERLs for: flouranthene, pyrene, phenanthrene, chrysene, BAP, baanthrene, anthracene, flourene, dibenzo(a,h)anthracene, total PCB, and pesticides; -Exceeded ERMs for PHEN, fluorene, ANTH, pyrene, BAA, BAP, HPAHs, and DDD at Station 18
Measure of Exposure: EqP Pore Water Toxic Unit	High	Medium	Low	Medium	> TU	adverse exposure	PHEN, FLUOR, ANTH, FLURAN, Pyrene and DDT exceeded TU

PORTSMOUTH NAVAL SHIPYARD
ASSESSMENT ENDPOINT INFORMATION SUMMARY

ASSESSMENT ENDPOINT: Benthic - Clark Cove							
MEASUREMENT ENDPOINT	DATA QUALITY	STRENGTH OF RELATIONSHIP	STUDY DESIGN	ENDPOINT WEIGHT	OUTCOME	INTERPRETATION	ISSUES
Measure of Effect Benthic Community Density	High	High	Medium	High	similar to reference, above 5th percentile	No effect	low density but high number of taxa (Sta. 3); Sta 5 &7 among the lowest densities in the study
Measure of Effect: Benthic Community Richness	High	High	Medium	High	similar to reference, above 5th percentile	No effect	see above
Measure of Effect Benthic Community Evenness	High	High	Medium	High	>95th percentile	Potential effect	Stations 4 and 8
Measure of Effect A. abdita toxicity	High	High	High	High	Not statistically significantly different from the control	No Effect	% survival ~100%
Measure of Exposure: Crustal ratio	High	Low	Medium	Medium	enriched	elevated exposure	Pb, Hg, Cr, Cd, As &Ag enriched at least > 3 times upper bound Highest concentration of As measured here
Measure of Exposure: Bulk Sediment Chemistry	High	Medium	Medium	Medium	> ERL for most COCs > ERM for tPCB, Hg, & pesticides	adverse exposure	-Exceeded NOAA ERLs for: flouranthene, pyrene, phenanthrene, baanthrene, anthracene, flourene, total PCB, DDT, Pb, Hg, Zn, Ni, Cr and Ag; -Exceeded ERMs for Hg, tPCB, ACHLOR, DDT, and DDD
Measure of Exposure: EqP Pore Water Toxic Unit	High	Medium	Medium	Medium	< TU	negligible exposure	no COCs exceeded TU
Measure of Exposure: SEM/AVS	High	High	Low	Medium	AVS >> SEM	neglibible exposure	no seasonal data on AVS and SEM

GLOSSARY OF TERMS

	Definition
Accuracy	The degree of agreement between a measured value and a true, expected value.
Acute Toxicity	The ability of a substance to cause effects resulting in severe biological harm within a short time after exposure to the toxic compound, usually within 24 to 96 hours.
Ag	Silver
Al	Aluminum
Algae	Microscopic plants which contain chlorophyll and live floating or suspended in water as **phytoplankton** in the **plankton**. They also may be attached to structures, rocks or other submerged surfaces. They are food for fish and small aquatic animals. **Algae** produce oxygen during sunlight hours and use oxygen during the night hours.
Ambient	Environmental or natural surrounding conditions.
ANTH	Anthracene - one of a number of **PAH** compounds.
Anthropogenic	Something made by humans, which effects nature.
AOC	Area of Concern - A specific contiguous area consisting of nearshore, intertidal, and subtidal habitats defined to evaluate whether ecological impacts are present. The Areas of Concern defined around the Portsmouth Naval Shipyard are listed below.
AOC - Back Channel (BC)	Nearshore and bottom sediment areas of the Back Channel adjacent to Seavey Island and extending from the back gate to the entrance of Jamaica Cove.
AOC - Clark Cove (CC)	The embayment formed by Seavey Island, Jamaica Island, Clark's Island, and the Clark's Island causeway and connecting to Portsmouth Harbor.
AOC - Dry Docks (DD)	The nearshore areas surrounding the dry docks and industrial areas on the western side of Seavey Island.
AOC - Jamaica Cove (JC)	The embayment formed by Seavey Island and Jamaica Island and connecting to the Back Channel.
AOC - Sullivan Point (SP)	The area located directly offshore of Sullivan Point and along the Piscataqua River side of the Clark's Island causeway.
AOC - DRMO Storage Yard (SY)	The area directly offshore of the Defense Reutilization and Marketing Office (DRMO) Storage Yard.
As	Arsenic
Assessment Endpoint	A component of the ecosystem that may be impacted by the stressors of concern, has ecological and societal value, and represents a component of the ecosystem that can be protected.
Avian Consumers	Birds of prey and water fowl (ducks, geese, gulls, and osprey) which feed on prey from the estuary.
AVS	Acid Volatile Sulfides - A reactive pool of sulfides that will bind with **divalent heavy metals** to form nontoxic and nonmobile compounds. These sulfides are released when sediments are treated with acid and the amount of sulfide released is referred to as **AVS** and the amount of metals that are simultaneously released is referred to as simultaneously extracted metal

Cord Grass	Grasses which commonly grow in intertidal **salt marshes**. The most prevalent species include ***Spartina alterniflora*** (tall and short varieties) in the low marsh and ***Spartina patens***, *Distichlis spicata*, and *Juncus gerardii* in the high marsh.
Cr	Chromium
Cu	Copper
DDD	dichlorodiphenyl dichloroethane - a metabolic breakdown product of **DDT**.
DDE	dichlorodiphenyl dichloroethene - a metabolic breakdown product of **DDT**.
DDT	dichlorodiphenyl trichloroethene - a pesticide compound, whose agriculture use in the U.S. was banned in 1973 because of its persistence in the environment and accumulation in the food chain.
Divalent	A chemical that can exist as an ion with a charge of 2^+ (e.g. Cd^{2+}, Cu^{2+}, Pb^{2+}, etc.).
Dose-Response	A quantitative relationship between the dose of a chemical and an effect caused by the chemical.
Dose-Response Curve	A graphical presentation of the relationship between degree of **exposure** to a chemical (dose) and observed biological effect or response.
EC_{20}	Effect Concentration 20% - the concentration of a chemical in air or water which is expected to cause an effect (other than death, e.g. reproductive impairment, reduced growth, biochemical response etc.) in 20% of test animals living in that air or water.
Ecological Receptors	Representative species selected to evaluate the likelihood of adverse impact to the **Assessment Endpoint**.
Ecosystem	An ecological system, a natural unit of living and nonliving components which interact to form a stable system in which a cyclic interchange of materials takes place between living and nonliving units.
Eelgrass	A **submerged aquatic plant** (*Zostera marina*) which can form meadows (eelgrass beds) that are capable of trapping sediment and providing habitat for a variety of birds, fish, and invertebrates.
Effects Assessment	The determination or estimation (qualitative or quantitative) of the magnitude, frequency, duration and extent of effects from exposure to a chemical.
Effects Measure	See **Measures of Effects.**
Effluent	Wastewater, treated or untreated, that flows out of a treatment plant, sewer, or industrial outfall into surface water.
Endpoint Weight	The ability (High, Medium, Low) of a measure to assess harm to the **Assessment Endpoint**. The **Endpoint Weight** was based on: Data Quality (attainment of data quality objectives), Strength of Association to the **Assessment Endpoint** (biological linkage, correlation of stressor to response, and utility for judging environmental harm), and Study Design (study design, site specificity, sensitivity of measurement, spatial representativeness, quantitativeness, and standardization of method). The overall endpoint weight was a qualitative average of the weights for Data Quality, Strength of Association, and Study Design.
Environmental Media	Components of the environment (water, **sediment**, and **biota**) that can accumulate contaminants.

	which are crystalline when solid, and many of which are characterized by opacity, ductility, conductivity, and a unique luster when freshly fractured. Metals will yield positively charged ions in aqueous solution of its salts.
Methylmercury	Any of several toxic compounds formed from metallic mercury by the action of microorganisms and capable of **bioaccumulating** in the **food chain**.
mg	Milligram - one-thousandth of a gram (0.000035 oz.).)
mg/L	Milligrams Per Liter - a measure of concentration of a dissolved substance. A concentration of one mg/L means that one milligram of a substance is dissolved in each liter of water which is equal to **parts per million (ppm)** since one liter of water is equal in weight to one million milligrams. For example: a liter of water containing 10 milligrams of calcium has 10 parts of calcium per one million parts of water, or 10 parts per million (10 ppm).
μg	Microgram - one-millionth of a gram (0.000000035 oz.).)
μg/L	Micrograms Per Liter - one microgram of a substance dissolved in each liter of water. This unit is equal to **parts per billion (ppb)** since one liter of water is equal in weight to one billion micrograms.
Mn	Manganese
Molecular Weight	The molecular weight of a compound in grams is the sum of the atomic weights of the elements in the compound.
Mortality	The proportion of deaths to population.
Ni	Nickel
NOAEL	No Observed Adverse Effect Level - the highest dose in an experiment which did not produce an observable adverse effect.
NOEL	No Observed Effect Level - in dose response experiments, the dose level at which no effects are noted.
Non-Point Source Pollution	Diffuse **pollution** sources that do not have a single point of origin or are not introduced into **receiving waters** from a specific outlet. The pollutants are generally carried off the land by stormwater runoff. The commonly used categories for non-point sources are agriculture. forestry, urban, mining, construction, dams and channels, land disposal, and saltwater intrusion.
Organic	Composed of plant or animal matter.
PAH	Polycyclic Aromatic Hydrocarbons - compounds containing more than one benzene ring in its structure.
Particulate	A very small solid suspended in water which can vary widely in size, shape, density, and electrical charge. **Colloidal** and dispersed particulates are artificially gathered together by the processes of coagulation and flocculation.
Partition Coefficient	A measure of the extent to which a chemical is divided between the soil/sediment and water phases.
Pb	Lead
PCB	Polychlorinated Biphenyl - any of several compounds that are produced by replacing hydrogen atoms in biphenyl with chlorine. Used in various industrial applications, tend to accumulate in animal tissues.
Pelagic Species	The community of organisms (fish, **plankton**) which spend the majority of their life floating or swimming in the water.

Risk Drivers	Chemicals that may be responsible for causing elevated risk.
Risk Management	The process for evaluating and selecting responses to risk.
Salt Marsh Community	Community of organisms living within a salt marsh (e.g. **cord grass** - ***Spartina***, snails, crabs, etc.).
SAV	Submerged Aquatic Vegetation - aquatic vegetation, such as sea grasses (eelgrass), that cannot withstand excessive drying and therefore live with their leaves at or below the water surface. SAVs provide an important habitat for young fish and other aquatic organisms.
Sediment	Matter which settles to the bottom in oceans, estuaries, rivers, lakes or other waterbodies.
SEM	Simultaneously Extracted Metal - the **heavy metals** associated with the reactive pool of **acid volatile sulfides**. The sulfides are released when sediments treated with acid and the amount of sulfide released is referred to as **AVS** and the amount of metals that are released simultaneously is referred to as **SEM**.
Spartina	The genus of commons species of **cord grasses** which grow in intertidal salt marshes. The main forms are *Spartina alterniflora* (tall and short varieties) and *Spartina patens*.
sumPCB	The sum of the measured PCBs.
Superfund	Federal law which authorizes EPA to manage the clean-up of abandoned or uncontrolled hazardous waste sites.
SWMU	Solid Waste Management Unit - an area designated in the Shipyard's Hazardous Waste Permit where hazardous materials may have been stored, treated, or released.
tDDx	Total **DDT** and metabolites (sum of **DDT**, **DDE**, and **DDD**).
Threshold	The lowest dose of a chemical at which a specified measurable effect is observed and below which it is not observed.
Toxic	A substance which is poisonous to an organism.
Toxic Pollutants	Materials contaminating the environment that cause death, disease. birth defects in organisms that ingest or absorb them. The quantities and length of exposure necessary to cause these effects can vary widely.
Toxic Substance	A chemical or mixture that may represent an unreasonable risk of injury to health or the environment.
Toxicant	A harmful substance or agent that may injure an exposed organism.
Toxicity	The quality or degree of being poisonous or harmful to plant, animal or human life.
Toxicity Assessment	Characterization of the toxicological properties and effects of a chemical, including all aspects of its absorption, metabolism, excretion and mechanism of action, with special emphasis on establishment of dose-response characteristics.
Toxicology	The science and study of poisons control.
tPCB	Total **PCB.**
Trophic Transfer	The process by which contaminants are accumulated in the **food chain**.
Turbidity	A measure of water cloudiness caused by suspended solids

1. EXECUTIVE SUMMARY

1.1 INTRODUCTION

Portsmouth Naval Shipyard is located on Seavey Island on the Piscataqua River between Maine and New Hampshire. In 1994, the Shipyard was added to the National Priorities List (NPL) which requires all cleanup activities to comply with the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA).

The USEPA conducted a Resource Conservation and Recovery Act (RCRA) Facility Assessment that identified several Solid Waste Management Units (SWMUs) where releases were known or suspected. Thirteen of the SWMUs evaluated by a RCRA Facility Investigation (RFI) and a Remedial Investigation and Feasibility Study (RI/FS). These areas included disposal areas, underground storage tanks (USTs), industrial waste outfalls, storage areas, and a 25-acre landfill located on the Shipyard property.

An Estuarine Ecological Risk Assessment (ERA) was conducted to evaluate risks to the estuary. The objectives of the ERA were:

- to assess the ecological risks to the offshore environments of the Piscataqua River and Great Bay Estuary from chemical stressors associated with the SWMUs at the Shipyard
- to develop information to support informed risk management decisions regarding remedial options
- to support communication of Shipyard associated ecological risks to the public

The purpose of this report is to evaluate ecological risks associated with the Shipyard's SWMUs. It is beyond the scope of the ERA to conduct a comparative analysis of all sources of risk in the estuary. This report also does not address human health concerns from the consumption of seafood or potential health effects from contact with water and sediment from the estuary. Risks to human health were addressed by the offshore human health risk assessment (McLaren/Hart 1994. Final Human Health Risk Assessment for Offshore Media for Portsmouth Naval Shipyard. Prepared for U.S. Department of Navy, Naval Facilities Engineering Command, Philadelphia, PA by McLaren/Hart Environmental Engineering Corporation, Albany, NY, May 1994.)

With several potential sources of environmental contamination in the lower Piscataqua River, monitoring stations were positioned to enhance detecting contamination originating from the Shipyard and evaluate the extent of the transport of contaminants released from the Shipyard. Stations were also selected to characterize the ecology of the lower Piscataqua River and its tributaries. Stations were located upstream, downstream, and cross-stream of the Shipyard and within Spruce Creek to provide information about the exposure and effect levels occurring in the lower estuary. Additional reference areas were sampled to provide information on ecological processes occurring outside of the influence of the lower Piscataqua River. Reference stations were selected in York River, ME, to provide measurements of ecological conditions in a nearby estuarine system with similar ecological characteristics, but without sources of industrial contamination. Stations were located in the upper Piscataqua River, Little Bay, and Great Bay to provide information on the potential far-field gradient of contaminants and identify potential sources of contamination upstream of the Shipyard. Reference stations were also established in relatively pristine areas to provide comparisons to regional background levels of contamination in mussel (Brave Boat Harbor), lobster (Isles of Shoals) and winter flounder (Gulf of Maine) tissues.

A weight-of-evidence approach was used to evaluate measures of effect and measures of exposure to interpret the level of risk evident for each applicable assessment endpoint and AOC. Since many of the assessment endpoints could not be measured directly, it was necessary to use measures of exposure and effect as indicators for the assessment endpoints. The measurement data had varying degrees of uncertainty, differing levels of reliability in terms of inferring harm to the assessment endpoint, and sometimes the measures produced conflicting results. Because no single measure alone is capable of determining whether there is risk or not, multiple lines of evidence were developed to characterize the magnitude of risk. Providing a systematic means of evaluating all the data, the weight-of-evidence analysis was used to derive the conclusions of risk and the confidence upon which the conclusion were based.

The weight-of-evidence approach evaluated whether the contaminants of concern (COCs) were responsible for potentially harmful environmental effects on the surrounding estuary. Using the weight-of-evidence approach enabled the multiple measures of exposure and effect to be related to the assessment endpoints. The method took into account the strengths and weaknesses of the measurement methods and evaluated how well the measures could assess harm to the assessment endpoints. To do this, each measurement of exposure and effect was assigned an endpoint weight (EW) of low, medium, or high. The endpoint weight served as a way to rank the relative uncertainty and reliability that was associated with the various measurements. The final overall endpoint weight was a qualitative average of three attribute scores: (1) Strength of Relationship, (2) Data Quality, and (3) Study Design. "Strength of Relationship" is the degree of association to the assessment endpoint, the relationship of the measure's response to the COCs, and the usefulness of the measure. "Data Quality" is an evaluation of whether the data met the data quality objectives defined for the measure. "Study Design" takes into account the site specificity, sensitivity of receptors, whether the area sampled is representative of possible site-related effects, the time of the year samples were taken, the quantity of data needed, and the use of a standard method (Menzie et al. 1996).

Ecological risk was determined by the correspondence between the exposure and effect data obtained for the assessment endpoints for each of the AOCs. Risk definitions were developed

1.3 RESULTS

Six AOCs were evaluated in the Estuarine ERA for the Shipyard: Clark Cove, Sullivan Point, DRMO storage yard, dry docks, Back Channel, Jamaica Cove (Figure 1-1). Using the weight of evidence approach, ecological risk was determined to exist in some of the AOCs (Table 1-3). There was:

- high confidence of intermediate risk from sediment exposure at Sullivan Point, the dry docks, and the Back Channel;
- high confidence of low risk from sediment exposure in Jamaica Cove;
- medium confidence of low risk from sediment and surface water exposure in Clark Cove;
- medium confidence of low risk from surface water exposure at Sullivan Point, at the dry docks, in the Back Channel, and in Jamaica Cove;
- medium confidence of negligible risk from surface water exposure at the DRMO storage yard; and
- medium confidence of negligible exposure to avian receptors in Portsmouth Harbor.
- No area of concern was identified as having high risk to any of the assessment endpoints evaluated.

Potential risk drivers with links to the SWMUs were identified for the following areas of concern:

- Clark Cove: seeps[1], chromium, nickel, polychlorinated-biphenyls
- Sullivan Point: seeps[2], copper, mercury, nickel, phenanthrene
- Dry Dock: copper, mercury, nickel, zinc, polychlorinated-biphenyls, phenanthrene, pyrene
- Back Channel: mercury, phenanthrene, fluorene, anthracene
- Jamaica Cove: seeps[3], lead

At least two compounds from the pesticide DDT (DDD, DDE, DDT, tDDx) were also identified as potential risk drivers in all the AOCs. The DDT compounds were not linked to any of the SWMUs. Although lead was fairly elevated in samples of biota, sediments, and seep samples collected around the Shipyard, lead did not appear to be a major risk driver. The levels of contamination measured were a mixture from a variety of sources and can not, with the data currently available, be attributed to specific origins.

---

[1] Because the seeps were not well characterized and they could be a direct route of release from some of the SWMUs, chemicals that exceeded chronic water quality criteria (WQC) in any of the seep samples were identified as potential risk drivers. Chemicals that exceeded water quality criteria in seep samples from Clark Cove were copper, nickel, mercury, and zinc.

[2] Chemicals that exceeded water quality criteria in seep samples from Sullivan Point were copper, mercury, and zinc.

[3] Chemicals that exceeded water quality criteria in seep samples from Jamaica Cove were copper, nickel, lead, and zinc.

Table 1-1 The route of exposure and measures of contaminants of concern (COC) concentrations that were conducted for the Estuarine ERA.

Route of Exposure	Measure of COC Concentration
Surface Water	Estuarine surface water Seep water Residues in deployed blue mussel
Sediment	Bulk sediment Acid volatile sulfide - simultaneously extracted metal
Biota	Residues in flounder liver and flesh Residues in lobster hepatopancreas and flesh Residues in native blue mussel Residues in eelgrass leaf and root Residues in fucoid algae Residues in spartina leaves

Table 1-2. The assessment endpoints, receptor species and measures of effect that were made for the Estuarine ERA.

Assessment Endpoint Receptor Species	**Measure**
Vitality of Pelagic Community	
Flounder	Flounder condition
Phytoplankton	Phytoplankton standing crop
Blue mussel	Water toxicity to deployed mussel growth
Sea urchin	Water toxicity to sea urchin gametes and larvae
Vitality of Epibenthic Community	
Lobster	Lobster abundance and condition
Fucoid algae	Fucoid algae abundance
Blue mussel	Mussel abundance and condition
Vitality of Infaunal Benthic Community	
Infaunal benthic community	Species richness, abundance, and evenness
Amphipods	Sediment toxicity to amphipods
Vitality of Eelgrass	
Eelgrass	Eelgrass abundance and growth
Vitality of Salt Marsh Community	
Cord grass/Salt Hay	Cord grass/Salt Hay abundance and growth
Salt marsh community	Distribution of plants and invertebrates
Vitality of Avian Community	
Black Duck	Tissue concentrations in prey species
Canada Goose	Tissue concentrations in prey species
Herring Gull	Tissue concentrations in prey species
Osprey	Tissue concentrations in prey species

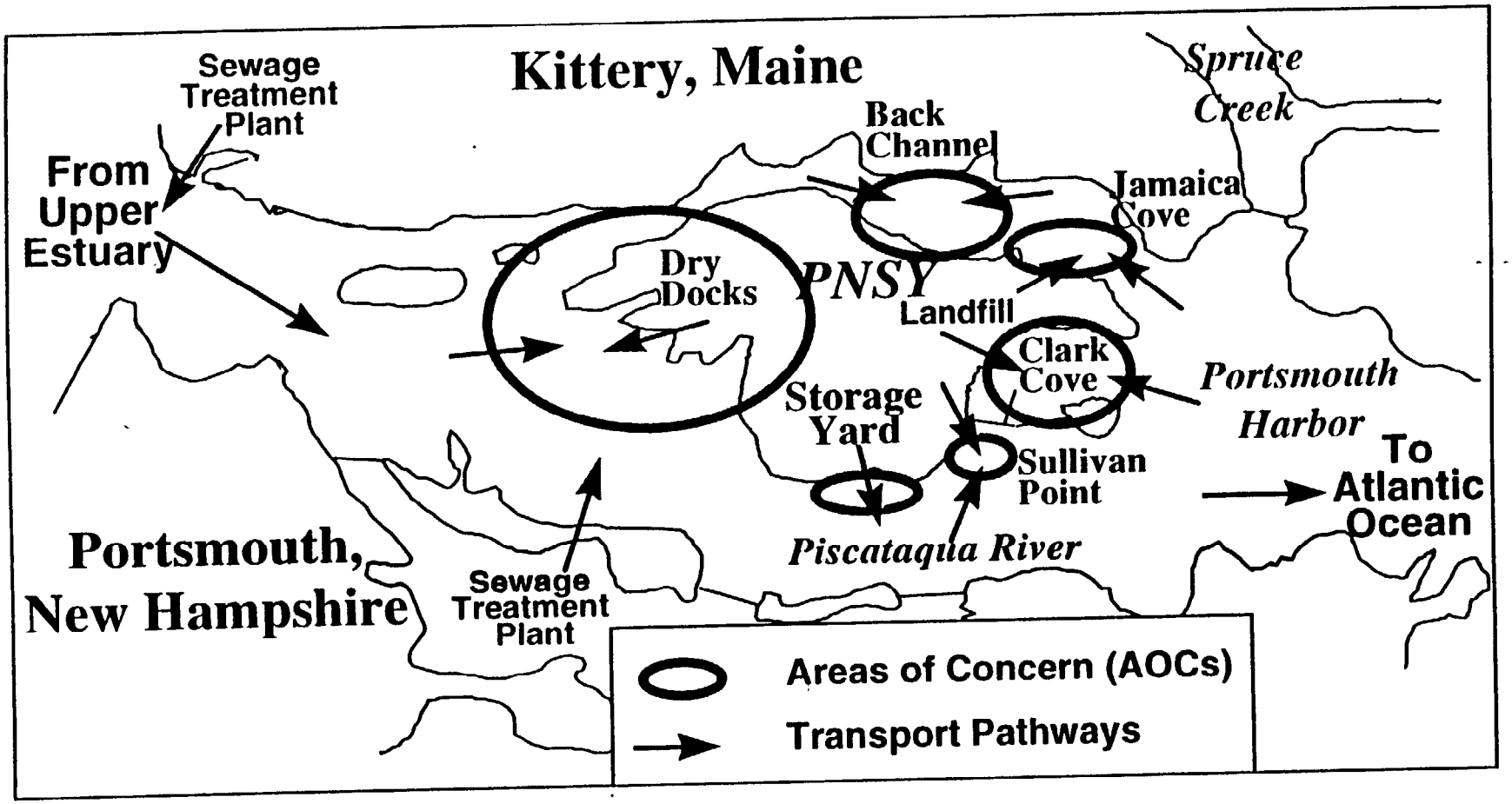


Figure 1-1. Diagram of the conceptual model developed for the Estuarine Ecological Risk Assessment.