6.0 WEIGHT-OF-EVIDENCE EVALUATION, UNCERTAINTY EVALUATION, AND CONCLUSIONS

The WOE approach was developed to integrate and evaluate the four lines of evidence collected at NFD Point Molate. The following section presents the WOE evaluation, evaluation of uncertainty, and conclusions.

The WOE evaluation includes the following:

- Summary of the WOE components completed as part of the ERA process (problem formulation, exposure and effects assessment, and risk characterization).
- Determination of finding, magnitude, and WOE score calculation.
- Presentation of WOE results for each NFD Point Molate sampling station.

The evaluation of uncertainty involves identifying sources associated with the ERA process that may potentially affect conclusions. Sources of uncertainty were identified through:

- Review of the ranks and weights assigned for measurement endpoints.
- Review of ERA assumptions.

The conclusions regarding risk to ecological resources at NFD Point Molate are based on the following:

- Site background and CSM.
- Analytical data collected in 1998.
- The assessment of exposure and effects.
- The characterization of risk presented.
- The integration of different lines of evidence collected for the offshore ERA.
- The evaluation of uncertainty.

6.1 SUMMARY OF WOE COMPONENTS COMPLETED AS PART OF THE ERA PROCESS

The WOE approach is integrated into both the Problem Formulation and Risk Characterization steps of the ERA. The following sections discuss how WOE components were used in each step.

6.1.1 WOE Components Considered in Problem Formulation

In the problem formulation step, assessment and measurement endpoints were selected, criteria were developed to evaluate measurement endpoint results for the purpose of defining the finding and magnitude of finding categories, and measurement endpoints were weighed according to procedures outlined in Section 2.8 and discussed in detail in Appendix B. The ranks used to calculate measurement endpoint weights, and the weights themselves, are shown in Table 2-5. The amphipod bioassay has the highest calculated weight (3.8), while the topsmelt and mysid bioassays have calculated weights of 3.44 and 3.38, respectively. The bioaccumulation endpoint has the next highest calculated weight (2.7), while the sediment chemistry endpoint has the lowest calculated weight (1.68). Measurement endpoint ranks and weight will be used to evaluate uncertainty in Section 6.4.

6.1.2 WOE Components Considered in Exposure and Effects Assessment

The WOE approach does not contribute to the exposure and effects assessments.

6.1.3 WOE Components Considered in Risk Characterization

In the risk characterization step, measurement endpoint results are evaluated by using finding criteria developed as part of the WOE approach to evaluate risk to the resources represented by the selected assessment endpoints.

6.2 DETERMINATION OF FINDINGS, MAGNITUDE, AND WOE SCORE CALCULATION

Based upon the criteria for findings and magnitudes outlined in Section 2.8, positive or negative findings were assigned, and the magnitude of each finding (i.e., high or low) was determined as shown in Table 6-1. The positive or negative magnitude for each measurement endpoint was then multiplied by the weight for that endpoint to derive a WOE score for each endpoint (as described in Section 2.8).

6.3 PRESENTATION OF WOE RESULTS FOR NFD POINT MOLATE SAMPLING STATIONS

The NFD Point Molate WOE ERA results are presented in two formats: (1) summary tables (Table 6-1), and (2) summary bar charts (Figure 6-1). Both the tables and bar charts are organized to group

measurement endpoint findings by the assessment endpoints they represent. The summary tables contain the information used to create the summary bar charts including the finding and magnitude associated with each measurement endpoint result, the calculated weight of each measurement endpoint, and the calculated WOE score. Appendix B provides further details on this approach.

The bar charts (Figure 6-1) depict measurement endpoint WOE scores with solid bars representing positive findings (i.e., positive indication of risk on a scale of 1 to 10) or negative findings (i.e., negative indication of risk on a scale of 1 to 10). As previously mentioned, the results are organized by assessment endpoint. This allows for the evaluation of risk to a particular assessment endpoint based on all of the measurement endpoint results associated with that assessment endpoint, or evaluation of sampling station risk based on the risk to all resources selected for protection (i.e., all assessment endpoints) (Figure 6-1).

As visually summarized in Figure 6-1, 9 of the 11 sampling stations evaluated for risk showed overwhelming correspondence between all of the assessment endpoints and their associated measurement endpoints, indicating no risk. For these sampling stations (DL-1, T11A, T2, T3-1, T5, T6, T9-1, T9-2 and P1-1) the WOE shows a negative indication of risk for all endpoints. For samplings locations T10-1 and T11-1, all endpoints show a negative indication of risk except for the sediment chemistry endpoint. Since the criteria used to develop findings for sediment chemistry were non-risk-based (i.e., comparison to ambient concentrations in the Bay), these findings in themselves cannot be used to indicate risk. When paired with the risk-based criteria (bioassay and bioaccumulation endpoints), the WOE for sampling stations T10-1 and T11-1 do not indicate risk.

6.4 EVALUATION OF UNCERTAINTY

The evaluation of uncertainty involves identifying sources of uncertainty associated with the ERA process that may potentially affect the conclusions of the assessment. According to USEPA (1996) *"Uncertainty analyses increase credibility by explicitly describing the magnitude and direction of uncertainties, and they provide that basis for efficient data collection of or application of refined methods."* Uncertainty associated with measurement endpoint results translates into uncertainty associated with the conclusions regarding risk at NFD Point Molate. To reduce the potential for uncertainty resulting in underestimates of actual risks at NFD Point Molate, conservative methods and procedures were used throughout the assessment.

The WOE process used to implement the offshore ERA at NFD Point Molate evaluates the strengths and weaknesses associated with various types of environmental data by considering important data attributes to weigh measurement endpoints (see Appendix B for more details). The measurement endpoint weight reflects the overall strength of the measurement endpoint, or the certainty that can be associated with its findings (e.g., how important a particular measurement endpoint finding is in determining potential risk). The data attributes directly evaluate the source of the certainty or uncertainty that can be associated with a measurement endpoint finding (e.g., identifies weaknesses of the measurement endpoint). Sources of uncertainty can be evaluated in a general sense by reviewing the ranks assigned to each data attribute of a measurement endpoint and the resulting weight for that measurement endpoint. Sources of uncertainty have been identified through a review of the ranks and weights assigned for NFD Point Molate measurement endpoints (Table 6-2). The WOE approach itself mitigates uncertainty associated with the offshore ERA as it takes the varying uncertainty associated with each measurement endpoint into account when integrating the measurement endpoint results to evaluate risk. Additional sources of uncertainty have been identified based on a review of the assumptions used to develop this ERA (e.g., assumptions associated with exposure parameter inputs to the dose equation and review of environmental data collected at NFD Point Molate). Identified sources of uncertainty are discussed in the following sections.

The discussion of uncertainty is organized by measurement endpoint. The relative uncertainty associated with each measurement endpoint has been determined based on the numerical weight calculated in the WOE for each measurement endpoint (see Table 2-5). The measurement endpoints will be evaluated in order of relative uncertainty with the endpoint having the highest level of uncertainty (sediment chemistry) being discussed first.

6.4.1 Uncertainty Associated With the Evaluation of Sediment Chemistry

The sediment chemistry measurement endpoint had the lowest calculated weight (1.68). This relatively low weight indicates that findings associated with this endpoint cannot be used to make judgments about potential risk with a high level of certainty. Sources of uncertainty can be identified by evaluating the individual ranks assigned to specific attributes. This measurement endpoint scored low for 6 of the 10 attributes including: biological linkage between measurement and assessment endpoint, correlation of stressor to response, utility of measure for judging environmental harm, site specificity, sensitivity of the measurement endpoint for detecting change, and spatial representativeness.

The primary sources of uncertainty for this endpoint are: (1) that the criteria used to evaluate the sediment chemistry data are not effects-based (e.g., exceedance of the criteria does not necessarily indicate the potential for deleterious effects), and (2) the criteria were not applied to representative samples (e.g., randomly sampled stations) from NFD Point Molate.

Effects-based criteria (e.g., ER-Ls and ER-Ms) are commonly used to evaluate sediment chemistry data and were considered for use at NFD Point Molate. Effects-based criteria are usually used as a conservative screening tool to indicate if there is a potential risk at a location or site requiring further evaluation. In cases where effects-based criteria indicate potential risk, further evaluation often consists of site-specific toxicological investigations (i.e., bioassays). The results of site-specific toxicological investigations are used to refine, and in some cases, re-define the risk estimate based on non-site-specific, generic, effects-based criteria. Comparisons of site data to effects-based criteria were not conducted at NFD Point Molate (i.e., not selected as a measurement endpoint) as multi-species site-specific bioassays were conducted to directly assess effects and potential risk. Instead, a non-effects-based criteria (a comparison to ambient conditions in San Francisco Bay) was used to determine if NFD Point Molate sediments should be considered contaminated relative to ambient.

To manage the uncertainty associated with this measurement endpoint, the following was conducted: (1) positive or negative findings were viewed only as corroborative information to evaluate results from other measurement endpoints (especially the bioassays), (2) the judgmental sampling approach was designed to provide a conservative estimate of risk by presenting "worst-case" locations, and (3) the sediment chemistry measurement endpoint contributed to a WOE evaluation of all of the data collected at each sampling location to evaluate overall risk.

6.4.2 Uncertainty Associated With the Evaluation of Benthic Invertebrate Bioaccumulation and Resulting Toxicity to Avian Receptors

A relatively low measurement endpoint weight (2.7) was calculated for the bioaccumulation measurement endpoint. Sources of uncertainty for this endpoint are associated with the following attributes: biological linkage between measurement and assessment endpoint, correlation of stressor to response, utility of measure for judging environmental harm and quantitative measure.

Many uncertainties potentially affect the results of comparing estimated doses of PAHs to shorebirds (based on bivalve body burdens) to an effects-based criteria associated with reproductive impairment

in birds. In general, the uncertainty can be associated with two sources: (1) the estimation of the dose to the shorebird, and (2) the derivation of the TRV for comparison to the dose.

6.4.2.1 Uncertainty Associated With the Dose Calculation

Because the parameters used to estimate dose (intake) are not always empirically measured, conservative assumptions were made which could result in an overestimate of exposure and risks. The dose calculations consider the concentration of the COPECs at an exposure point, physical characteristics of the receptor, and the exposure frequency. Each of these three inputs has varying degrees of uncertainty.

The concentrations of COPECs at the exposure point were directly measured (i.e. site-specific bivalve tissue and sediment data were collected). Therefore, food-chain analysis and modeling, which are typically sources of uncertainty, were not conducted. However, there were uncertainties associated with the assumptions made about the percent diet composition and foraging behavior of the receptors. The use of bivalves to monitor environmental concentrations of PAHs is considered to be conservative as bivalves are limited in their ability to metabolize these compounds and, therefore, are representative of maximum bioconcentration. This is the principle behind Mussel Watch Programs which provide an indication of water quality at steady-state conditions in a given locality. In the risk assessment paradigm, clam tissue burdens are representative of the "worst case" scenario of contaminant exposure. It follows then, that if the calculated risk from these accumulated burdens (either to the clam or to potential consumers) is found to be minimal, there should be no excess risk to sensitive in situ populations. As a result, the worst case is conservative and hence protective. The diet for the birds considered at NFD Point Molate can be highly variable (e.g., mullusks, polychaete worms, crustaceans, and vegetation) and is dependent on specific prey availability. The assumption that the birds' diet consists solely of bivalves is, therefore, likely an over-estimate of exposure.

Foraging habits that impact the amount of incidental sediment ingested also directly affect exposure estimates. The percent sediment ingestion for the western sandpiper was taken from the literature (18 percent of diet), but no reported value was available for the scaup. Therefore, a conservative value for the scaup was estimated (11 percent of diet) using the percent sediment ingestion for a surrogate species with similar feeding habits (i.e., the wood duck). The natural history information gathered for the western sandpiper and the scaup indicate that both estimates for percent diet composition of prey and sediment are based on very conservative assumptions (see Appendix H for natural history information).

Physical characteristics of the receptors that affect the dose calculation include body weight and daily ingestion rate. A mean body weight for the western sandpiper was derived by taking a mean from available literature for both sexes. However, the scaup weight was more variable because both the greater and lesser scaup were considered and because less body weight data were available. Body weights vary widely, and lower body weights are associated with higher calculated doses (when the ingestion rate was derived from an allometric equation). Because a low body weight was chosen (by using the mean of female lesser scaup weights available in the literature), this is likely to provide a conservative estimate of exposure to the population as a whole.

Daily ingestion rates were not available in the literature for either species, so an ingestion rate was estimated using an allometric equation developed by Nagy (1987) which is based on body weight. The allometric equation is based on the assumption that as body weight increases, ingestion rate would also increase by a constant rate. There is a large amount of uncertainty in estimating ingestion based on an allometric equation. However, the equation is designed to be conservative and would likely overestimate ingestion rates.

The exposure frequency for each receptor is based on the amount of time the species uses the site. Receptors used in the risk analysis for NFD Point Molate were assumed to be year-round residents and have ranges such that they forage and live within the offshore area of NFD Point Molate 100 percent of the time (i.e., a site use factor of 1 was used in the dose calculation). The natural history information for the western sandpiper and the scaup reveal that both species are migratory and spend several months a year migrating to and from their breeding territories in Alaska or Canada. The western sandpiper exhibits high site fidelity in its wintering territory, but its estimated wintering home range is far greater than the offshore area at Point Molate. The scaup does not have a wintering home range as such, rather it is thought to move from site to site based on habitat and prey availability. The time spent at the NFD Point Molate for the scaup is, therefore, likely to be far less than 100 percent of its time, however no literature is available to confirm this. Thus, the assumption that either species would use the offshore environment at NFD Point Molate 100 percent of the time is a conservative assumption. The assumption is made even more conservative by the fact that risk is evaluated at each individual sample location. Thus, the SUF of 1 actually assumes that the bird forages 100 percent of the time within the area of any given sample location. Since the sample locations were chosen to target worst case scenarios, the SUF of 1 is likely to result in a large overestimation of dose.

6.4.2.2 Uncertainty Associated With the Development of Criteria to Evaluate Estimated Doses (TRVs)

A formal avian TRV for sum PAH could not be developed; rather ranges were developed in which high and low magnitudes for both positive and negative risk findings were defined. Because of the variability and uncertainty associated with the derivation of these criteria, an uncertain risk range was also defined. made in the development of the ranges. These include: (1) all toxicity associated with crude oil was assumed to be caused by sum PAH; (2) for risk determination boundaries, the amount of PAH in crude oil was assumed to be the lowest value for any crude composition found in the literature; and (3) the reproductive endpoints demonstrated in the literature were assumed to be the result of the actual crude oil exposure rather than the result of confounding factors within the study (such as decreased food consumption).

No avian studies were available for specific PAHs or sum PAH. Therefore, studies on the effects of crude oils and whole petroleum products on avian species were reviewed. The assumption was made that all of the toxicity associated with the oil was caused by the sum PAH. While this is not likely to be the case, this assumption ensures conservatism by assuming that a relatively small PAH concentration in crude oil is solely responsible for any observed effect (i.e., all toxicity) in the studies reviewed. In order to calculate a sum PAH dose for each study, a sum PAH number for each of the crude oils was derived. In order to calculate a sum PAH dose for each study, it was necessary to derive a sum PAH number for each of the crude oils. Compositional data for the crudes used in the studies were not available; therefore, a range was derived based on several other crude oils with known composition. This yielded a range for sum PAH in crude oil (0.34 to 2.1 percent), creating variability within the dose ranges and thus within the findings. Because the lowest PAH composition required consideration, the dose for a low magnitude positive finding varied almost an order of magnitude (0.5 to 9 mg/kg/d). As such, the criteria reflected by a toxicity figure developed for avian TRVs are likely to be extremely conservative for risk estimation.

6.5 UNCERTAINTY ASSOCIATED WITH LABORATORY TOXICITY TESTS

All three laboratory toxicity tests (the amphipod bulk sediment bioassay, mysid SWI bioassay, and topsmelt SWI bioassay) conducted to evaluate sediment toxicity at NFD Point Molate have high calculated weights (3.44 to 3.80) relative to the sediment chemistry and bioaccumulation endpoints. The bulk sediment bioassay had the highest calculated weight (3.80) of the three bioassays, which reflects a lower level of uncertainty than is associated with the two SWI bioassays. In addition to having a different level of uncertainty, the bulk sediment bioassay has different sources of uncertainty than the two SWI

bioassays. Based on the ranks assigned to the bulk sediment bioassay for measurement attributes, sources of uncertainty are potentially associated with the sensitivity of the measurement endpoint, the degree of association between the measurement and assessment endpoint and the stressor response relationship evaluated by the measurement endpoint. Based on ranks assigned for measurement endpoint attributes for the two SWI tests, the most significant sources of uncertainty are associated with the WOE attributes, sensitivity of the measurement endpoint for detecting change, and standard measure. The uncertainties, and sources of uncertainty associated with bioassays conducted at NFD Point Molate, are discussed in more detail in the following sections.

6.5.1 Uncertainty Associated With Using SWI Bioassay Test Systems

Uncertainty associated with using SWI bioassay test systems is discussed in the following two sections: (1) standard procedures and evaluation criteria, and (2) the sensitivity of the SWI bioassays.

6.5.1.1 Standard Procedures and Evaluation Criteria

SWI bioassay test systems have not been routinely used; therefore, standard procedures and environmentally relevant criteria to judge them have not been developed. As part of the problem formulation process, it was determined that a MSD approach would be used to evaluate SWI test results for NFD Point Molate. This approach is often used to evaluate aquatic and marine bioassay results.

The MSD is a value that indicates the difference between reference and test site mean toxicity that will be considered statistically significant given the inherent level of replicate variation in reference tests (SFEI, 1996).

MSDs have been calculated for many toxicity test protocols and may require 50 to 100 independent tests in order for them to have any biological relevance. SWI tests are very new and, in the case of these two species, are still under development. Therefore, rigorously defined MSDs for these taxa were not available. In lieu of species-specific MSDs, a generic MSD was considered. To manage the uncertainty of using a generic MSD as an effects-based criteria to evaluate site-specific toxicological data, a conservative MSD value was selected. As a generic MSD, 20 percent was selected to evaluate SWI test results. This value is considered to be conservative, as the only MSD developed for a SWI bioassay is a MSD of 41 percent which has been calculated by the DFG's MPSL for the purple sea urchin larval development SWI bioassay. It is not surprising that MSDs for SWI tests are higher than other types of

bioassay protocols since the SWI protocol specifies the use of intact sediment cores for each replicate of the bioassay. In contrast, other sediment tests use homogenized sediments for replicate tests. Spatial variation among intact sediment samples may be expected to contribute a greater amount of variation to toxicity compared to tests using homogenized sediment.

6.5.1.2 Sensitivity of SWI Bioassay

The sensitivity of the SWI bioassay species used (topsmelt and mysid) at NFD Point Molate is unknown. However, both test organisms have been established as sensitive water-column test species relative to other typical bioassay test species using standard laboratory reference toxicant tests. SWI testing involves the evaluation of potentially contaminated sediments that represent complex matrices and may contain mixtures of contaminants. Many factors affect the mobility and bioavailability of contaminants associated with sediments (e.g., similarly contaminated sediments may exhibit highly variable toxicity based on many factors). Therefore, a SWI result cannot be directly associated with a sediment concentration of a constituent. Although the sensitivities of the SWI bioassay conducted (mysid and topsmelt) are not known, both species are expected to be sensitive relative to other standard test species.

6.5.2 Specific Uncertainty Associated With the *M. bahia* SWI Test

Use of a Surrogate Species

M. bahia does not occur at NFD Point Molate. *M. bahia* was selected as a surrogate species because of its sensitivity and applicability. There is an inherent level of uncertainty in using a surrogate species to conduct site-specific toxicological evaluations. The uncertainty associated with using surrogate species is managed by selecting conservative criteria for evaluation, appropriate species, and established protocol. *M. bahia* is a standard, water column test species which has been extensively used in the evaluation of drilling muds and dredge sediments and is generally recognized as one of the most sensitive water column invertebrate test species having a growth and survival endpoint.

Uncertainties associated with the *M. bahia* measurement endpoint are mitigated primarily through: (1) the use of a second bioassay (an amphipod bulk sediment bioassay) to evaluate toxicity to the benthic invertebrate community, and (2), by considering a second anecdotal toxicity endpoint for the mysid (growth). The bulk sediment bioassay using the amphipod *E. estuarius* is a standard protocol with well developed evaluation criteria. No effects were observed in the bulk sediment bioassay: this result corroborates the mysid SWI results indicating no effects. Growth was evaluated as an anecdotal endpoint

for the mysid SWI bioassays to evaluate potential effects. Mean/mysid weights for laboratory replicates at all NFD Point Molate sampling locations (0.22 mg to 0.29 mg) were equal to or greater than the mean weight for the laboratory control (0.22 mg) providing further indication of no effects.

In summary, the agreement between the bulk sediment bioassay results and the mysid growth survival results minimize uncertainties associated with the mysid SWI bioassay measurement endpoint.

6.5.3 Uncertainty Associated With the *A. affinis* SWI Test

The primary source of uncertainty associated with the topsmelt SWI test is associated with the performance of the reference location. NFD Point Molate sample location results from this bioassay are compared to mean survival in the reference location (Paradise Cove). Comparison of hatchability results for this bioassay to the reference location using the MSD approach indicate no effects would be expected greater than ambient. However, mean survival for five laboratory replicates at Paradise Cove was 76 percent, which indicates some ambient toxicity at the reference site. Thus, the performance of the reference location is a source of uncertainty.

The uncertainty associated with this endpoint can be mitigated in two ways. First, the reference location could be assumed to have a higher survival consistent with no toxicity. For example, if the reference location was assumed to have 90 percent hatchability (a result consistent with an acceptable laboratory control) and the NFD Point Molate data were compared to that theoretical result using the MSD approach, the uncertainty associated with the reference location would become irrelevant. When the NFD Point Molate sample location results (74 to 100 percent hatchability) are compared to the 20 percent MSD criteria associated with a theoretical reference result of 90 percent (e.g., >70 percent hatchability), no effects (i.e., positive findings) are predicted. This theoretical evaluation of the data for this measurement endpoint indicates that the uncertainty associated with the performance of the reference location is minimal and the results of this measurement should be considered highly applicable.

6.5.4 Uncertainty Associated With the Bulk Sediment Bioassay Using *E. estuarius*

The bulk sediment bioassay using *E. estuarius* had the highest weight (3.80) of all the measurement endpoints and, therefore, the highest level of certainty can be associated with its results and the conclusions drawn based on those results. There are still, however, uncertainties associated with the application of this assay. *E. estuarius*, a test species that does not live in San Francisco Bay, is used as a

surrogate species to evaluate the toxicity of Bay sediments to associated benthos. Thus, there is inherent uncertainty associated using this bioassay to evaluate the potential for actual impacts to San Francisco Bay benthos.

The uncertainties associated with using a surrogate species to evaluate the toxicity of San Francisco Bay sediments is mitigated by the fact at *E. estuarius* has been extensively used in San Francisco Bay and highly applicable criteria have been developed to evaluate *E. estuarius* toxicity tests. *E. estuarius* is a burrowing amphipod which is found in fine intertidal sediments from British Columbia to Central California (Hoffman et al., 1995). *E. estuarius* has and continues to be used by the San Francisco Estuary Institute's Regional Monitoring Program for Trace Substances and was one of the species used by the CRWCB to develop the report "Evaluation and Use of Sediment Reference Sites and Toxicity Tests in San Francisco Bay" (CRWCB, 1998). *E. estuarius* was found to rank well with respect to test success rate, variability, tolerance to confounding factors such as grain size, and ability to distinguish between sediments from impacted and reference sites (CWRCB, 1998).

The uncertainty associated with using *E. estuarius* to evaluate the potential risk to the benthic invertebrate community at NFD Point Molate is additionally mitigated by the fact that a second bioassay was conducted (the mysid SWI test) and compared to the *E. estuarius* results. Both bioassays indicated no effects; therefore, it can be concluded that there is no indication of risk to the benthic invertebrate community at NFD Point Molate.

6.6 UNCERTAINTY ASSOCIATED WITH THE LABORATORY EXPOSURE MODEL (IN SITU VS. LABORATORY ENVIRONMENTAL CONDITIONS)

Sediment samples collected at NFD Point Molate for toxicological evaluation was collected as bulk sediment samples, and SWI samples. Bulk sediment samples are collected in a manner that does not preserve the SWI, and the overlying water is decoupled from the sediment at the time of collection. SWI samples are collected in a manner that preserves the SWI and retains the overlying water at the time the sample is collected. Both samples, however, are tested in the laboratory using different overlying water (e.g., clean laboratory control water) and different environmental conditions (e.g., temperature and salinity) than at the location of collection.

Therefore, uncertainties associated with the replacement of site water with laboratory water to evaluate the sediments for toxicity. Studies have shown that changes in overlying water quality characteristics (i.e., pH, temperature, and salinity) can mobilize sediment-bound chemicals (Forstner, 1987). Studies

have indicated that changes in salinity and pH can affect the mobility of chemicals from sediments into the overlying water (Salomons et al., 1987), thus affecting sediment toxicity. Temperature changes can also have a marked effect on toxicity, but no single pattern for effects due to temperature changes are known (Rand, 1995).

As reflected in the data presented in Table 6-3, there are differences in salinity and temperature between the laboratory and the field. Data were not collected for pH at NFD Point Molate sampling locations, however, the values collected in the laboratory are consistent with what would be expected for a marine system such as the offshore environment at NFD Point Molate. The overall effect of replacing the overlying site water with laboratory water in the present investigation is unclear, but given that the temperature and salinity ranges measured in the field and laboratory were similar, and the pH is expected to be similar, effects on toxicity are expected to be minimal.

6.7 CONCLUSIONS FOR WOE AND UNCERTAINTY EVALUATION

The assessment endpoint, protection of the benthic invertebrate community associated with NFD Point Molate sediments, was evaluated using three measurement endpoints: (1) a bulk sediment amphipod bioassay, (2) a mysid SWI bioassay, and (3) sediment chemistry. No effects were observed in either of the two bioassay endpoints. The sediment chemistry endpoint is not effects-based, therefore, effects were not evaluated for this endpoint. Results from both bioassays indicate that there is no risk to the benthic community. Based on the results of the uncertainty analysis, the highest relative level of certainty was associated with the bulk sediment bioassay, followed by the mysid bioassay and the sediment chemistry measurement endpoint. Sources of uncertainty have been identified and managed, and the level of uncertainty associated with the measurement endpoint results was considered to be acceptable for risk evaluation purposes. The WOE approach integrated all of the measurement endpoint results while considering uncertainty to make conclusions about potential risk. The WOE evaluation indicates that there is no risk to the benthic invertebrate community by potentially contaminated sediments at NFD Point Molate. Furthermore, based on the evaluation of uncertainty, these results are sufficient to make risk management decisions. No further data collection is recommended for this assessment endpoint.

The assessment endpoint, protection of the larval fish community associated with NFD Point Molate eelgrass beds, was evaluated using two measurement endpoints: (1) a topsmelt SWI bioassay, and (2) sediment chemistry. No effects were observed in the bioassay endpoint. The sediment chemistry endpoint was not effects-based, therefore, effects were not evaluated for this endpoint. Bioassay results

indicate that there is no risk to the larval fish community at NFD Point Molate. The highest relative level of certainty was associated with the mysid SWI bioassay, followed by the sediment chemistry measurement endpoint. Sources of uncertainty have been identified and managed and the level of uncertainty associated with the measurement endpoint results was considered to be acceptable for risk evaluation purposes. The WOE evaluation indicates that there is no risk to the larval fish community at NFD Point Molate. Based on the evaluation of uncertainty, these results are sufficient to make risk management decisions. No further data collection is recommended for this assessment endpoint.

The assessment endpoint, protection of the shorebird community that forages in the intertidal mudflats at NFD Point Molate, was evaluated using two measurement endpoints: (1) a comparison of calculated sum PAH doses to two avian receptors to an effects-based criterion, and (2) sediment chemistry.

Sum PAH doses calculated for the scaup at intertidal NFD Point Molate sampling stations (i.e., DL-1, T3-1, T5, T6, T9-1 and T9-2) are representative of no adverse effects to avian receptors. Doses calcu-lated for the remaining intertidal sites (T-2, T11-1, T10-1 and T11A) fell in the undetermined category (between the criteria defining negative and positive findings). The potential risk is undetermined at these sites. All of the doses calculated for the western sandpiper fell between the criteria defining negative and positive findings (i.e., in the undetermined category). The potential risk to avian receptors is indeterminate based on doses modeled to the western sandpiper. The sediment chemistry endpoint is not effects-based, therefore, effects were not evaluated for this endpoint. The WOE evaluation indicates that there is no risk to the shorebird community. The evaluation of uncertainty also identified that the parameters used to estimate doses were highly conservative. The WOE conclusions, teamed with the conservative assumptions identified in the uncertainty analysis, indicate that there is no risk to the shorebird community from exposure to NFD Point Molate sediments. These results are considered to be sufficient to make risk management decisions. No further data collection is recommended for this assessment endpoint.

Ecological endpoints (i.e., assessment and measurement endpoints) are explicit statements which identify desired environmental goals and provide a means for determining whether an unacceptable effect may occur. The assessment and measurement endpoints for NFD Point Molate represent those ecological resources in the intertidal habitat at the site selected for protection. The WOE approach allowed for an integration of the multiple lines of evidence (e.g., measurement endpoints) collected for each assessment endpoint at NFD Point Molate by:

- Evaluating the strength of the measurement endpoints in predicting risk to the ecological resources at NFD Point Molate recognized for protection.

- Integrating those lines of evidence, considering uncertainty, to formulate conclusions regarding risk.

Table 6-4 summarizes the conclusions regarding risk to each of the three NFD Point Molate assessment endpoints. In two of the assessment endpoints (benthic invertebrates and larval fish), the WOE indicates that there is no risk from offshore sediments at NFD Point Molate to benthic invertebrates and larval fish. The WOE indicates indeterminate risk for shorebirds. However, the conservatism used in evaluating this assessment endpoint makes it likely that an undetermined finding would become a negative finding using more realistic exposure parameters (e.g., site use factor dietary composition for the representative avian species).

Based on the evaluation presented in this report, the following overall conclusions can be made:

- The assessment endpoints represent three ecological resources at NFD Point Molate.
- The measurement endpoint results indicate that there is no risk to the resources selected for protection (i.e., assessment endpoints) at NFD Point Molate.
- No further data needs are identified; thus, these results are considered to be sufficient to make risk management decisions.
- There is no risk from offshore sediments at NFD Point Molate to ecological resources.

TABLE 6-1

SUMMARY TABLE WITH FINDINGS, MAGNITUDE OF FINDINGS, WEIGHTS, WOE SCORES AND WOE RANKING SCORES

Sample Station DL-1	Finding	Finding & Magnitude	Weight	WOE Score	
Measurement Endpoint	positive (+) negative (-) or undetermined (u)	Finding - + or - Magnitude - 1 = low and 2 = high	(1-5)	Magnitude x Weight	
A1-Bulk Sediment Bioassay	-	2	3.80	7.60	8
A3-SWI (Mysid)	-	2	3.38	6.76	7
C-Chemistry Screening	-	1	1.68	1.68	2
A2-SWI (Topsmelt)	-	2	3.44	6.88	7
C-Chemistry Screening	-	1	1.68	1.68	2
B1-Bioaccumulation (Sandpiper)	u	0	2.70	0.00	0
B2-Bioaccumulation (Scaup)	-	1	2.70	2.70	3
C-Chemistry Screening	-	1	1.68	1.68	2

Sample Station T10-1	Finding	Finding & Magnitude	Weight	WOE Score	
Measurement Endpoint	positive (+) negative (-) or undetermined (u)	Finding - + or - Magnitude - 1 = low and 2 = high	(1-5)	Magnitude x Weight	
A1-Bulk Sediment Bioassay	-	2	3.80	7.60	8
A3-SWI (Mysid)	-	2	3.38	6.76	7
C-Chemistry Screening	+	1	1.68	1.68	2
A2-SWI (Topsmelt)[1]	-	2	3.44	6.88	7
C-Chemistry Screening	+	1	1.68	1.68	2
B1-Bioaccumulation (Sandpiper)	u	0	2.70	0.00	0
B2-Bioaccumulation (Scaup)	u	0	2.70	0.00	0
C-Chemistry Screening	+	1	1.68	1.68	2

Sample Station T11A	Finding	Finding & Magnitude	Weight	WOE Score	
Measurement Endpoint	positive (+) negative (-) or undetermined (u)	Finding - + or - Magnitude - 1 = low and 2 = high	(1-5)	Magnitude x Weight	
A1-Bulk Sediment Bioassay	-	2	3.80	7.60	8
A3-SWI (Mysid)	-	2	3.38	6.76	7
C-Chemistry Screening	-	1	1.68	1.68	2
A2-SWI (Topsmelt)[1]	-	2	3.44	6.88	7
C-Chemistry Screening	-	1	1.68	1.68	2
B1-Bioaccumulation (Sandpiper)	u	0	2.70	0.00	0
B2-Bioaccumulation (Scaup)	u	0	2.70	0.00	0
C-Chemistry Screening	-	1	1.68	1.68	2

TABLE 6-1
(Continued)

SUMMARY TABLE WITH FINDINGS, MAGNITUDE OF FINDINGS, WEIGHTS, WOE SCORES AND WOE RANKING SCORES

Sample Station T11-1	Finding	Finding & Magnitude	Weight	WOE Score	
Measurement Endpoint	positive (+) negative (-) or undetermined (u)	Finding - + or - Magnitude - 1 = low and 2 = high	(1-5)	Magnitude x Weight	
A1-Bulk Sediment Bioassay	-	1	3.80	3.80	4
A3-SWI (Mysid)	-	2	3.38	6.76	7
C-Chemistry Screening	+	1	1.68	1.68	2
A2-SWI (Topsmelt)	-	2	3.44	6.88	7
C-Chemistry Screening	+	1	1.68	1.68	2
B1-Bioaccumulation (Sandpiper)	u	0	2.70	0.00	0
B2-Bioaccumulation (Scaup)	u	0	2.70	0.00	0
C-Chemistry Screening	+	1	1.68	1.68	2

Sample Station T2	Finding	Finding & Magnitude	Weight	WOE Score	
Measurement Endpoint	positive (+) negative (-) or undetermined (u)	Finding - + or - Magnitude - 1 = low and 2 = high	(1-5)	Magnitude x Weight	
A1-Bulk Sediment Bioassay	-	1	3.80	3.80	4
A3-SWI (Mysid)	-	2	3.38	6.76	7
C-Chemistry Screening	-	1	1.68	1.68	2
A2-SWI (Topsmelt)	-	2	3.44	6.88	7
C-Chemistry Screening	-	1	1.68	1.68	2
B1-Bioaccumulation (Sandpiper)	u	0	2.70	0.00	0
B2-Bioaccumulation (Scaup)	u	0	2.70	0.00	0
C-Chemistry Screening	-	1	1.68	1.68	2

Sample Station T3-1-1	Finding	Finding & Magnitude	Weight	WOE Score	
Measurement Endpoint	positive (+) negative (-) or undetermined (u)	Finding - + or - Magnitude - 1 = low and 2 = high	(1-5)	Magnitude x Weight	
A1-Bulk Sediment Bioassay	-	1	3.80	3.80	4
A3-SWI (Mysid)	-	2	3.38	6.76	7
C-Chemistry Screening	-	2	1.68	3.36	3
A2-SWI (Topsmelt)	-	2	3.44	6.88	7
C-Chemistry Screening	-	2	1.68	3.36	3
B1-Bioaccumulation (Sandpiper)	u	0	2.70	0.00	0
B2-Bioaccumulation (Scaup)	-	1	2.70	2.70	3
C-Chemistry Screening	-	2	1.68	3.36	3

TABLE 6-1
(Continued)

SUMMARY TABLE WITH FINDINGS, MAGNITUDE OF FINDINGS, WEIGHTS, WOE SCORES AND WOE RANKING SCORES

Sampling Station T5	Finding	Finding & Magnitude	Weight	WOE Score	
Measurement Endpoint	positive (+) negative (-) or undetermined (u)	Finding - + or - Magnitude - 1 = low and 2 = high	(1-5)	Magnitude x Weight	
A1-Bulk Sediment Bioassay	-	2	3.80	7.60	8
A3-SWI (Mysid)	-	2	3.38	6.76	7
C-Chemistry Screening	-	2	1.68	3.36	3
A2-SWI (Topsmelt)	-	2	3.44	6.88	7
C-Chemistry Screening	-	2	1.68	3.36	3
B1-Bioaccumulation (Sandpiper)	u	0	2.70	0.00	0
B2-Bioaccumulation (Scaup)	-	1	2.70	2.70	3
C-Chemistry Screening	-	2	1.68	3.36	3

Sample Station T6	Finding	Finding & Magnitude	Weight	WOE Score	
Measurement Endpoint	positive (+) negative (-) or undetermined (u)	Finding - + or - Magnitude - 1 = low and 2 = high	(1-5)	Magnitude x Weight	
A1-Bulk Sediment Bioassay	-	1	3.80	3.80	4
A3-SWI (Mysid)	-	2	3.38	6.76	7
C-Chemistry Screening	-	2	1.68	3.36	3
A2-SWI (Topsmelt)	-	2	3.44	6.88	7
C-Chemistry Screening	-	2	1.68	3.36	3
B1-Bioaccumulation (Sandpiper)	u	0	2.70	0.00	0
B2-Bioaccumulation (Scaup)	-	1	2.70	2.70	3
C-Chemistry Screening	-	2	1.68	3.36	3

Sample Station T9-1	Finding	Finding & Magnitude	Weight	WOE Score	
Measurement Endpoint	positive (+) negative (-) or undetermined (u)	Finding - + or - Magnitude - 1 = low and 2 = high	(1-5)	Magnitude x Weight	
A1-Bulk Sediment Bioassay	-	1	3.80	3.80	4
A3-SWI (Mysid)	-	2	3.38	6.76	7
C-Chemistry Screening	-	1	1.68	1.68	2
A2-SWI (Topsmelt)	-	2	3.44	6.88	7
C-Chemistry Screening	-	1	1.68	1.68	2
B1-Bioaccumulation (Sandpiper)	u	0	2.70	0.00	0
B2-Bioaccumulation (Scaup)	-	1	2.70	2.70	3
C-Chemistry Screening	-	1	1.68	1.68	2

TABLE 6-1
(Continued)

SUMMARY TABLE WITH FINDINGS, MAGNITUDE OF FINDINGS, WEIGHTS, WOE SCORES AND WOE RANKING SCORES

Sample Station T9-2	Finding	Finding & Magnitude	Weight	WOE Score	
Measurement Endpoint	positive (+) negative (-) or undetermined (u)	Finding - + or - Magnitude - 1 = low and 2 = high	(1-5)	Magnitude x Weight	
A1-Bulk Sediment Bioassay	-	1	3.80	3.80	4
A3-SWI (Mysid)	-	2	3.38	6.76	7
C-Chemistry Screening	-	1	1.68	1.68	2
A2-SWI (Topsmelt)	-	2	3.44	6.88	7
C-Chemistry Screening	-	1	1.68	1.68	2
B1-Bioaccumulation (Sandpiper)	u	0	2.70	0.00	0
B2-Bioaccumulation (Scaup)	-	1	2.70	2.70	3
C-Chemistry Screening	-	1	1.68	1.68	2

Sampling Station P1-1	Finding	Finding & Magnitude	Weight	WOE Score	
Measurement Endpoint	positive (+) negative (-) or undetermined (u)	Finding - + or - Magnitude - 1 = low and 2 = high	(1-5)	Magnitude x Weight	
A1-Bulk Sediment Bioassay	-	1	3.80	3.80	4
A3-SWI (Mysid)	-	2	3.38	6.76	7
C-Chemistry Screening	-	1	1.68	1.68	2
A2-SWI (Topsmelt)	-	2	3.44	6.88	7
C-Chemistry Screening	-	1	1.68	1.68	2
B1-Bioaccumulation(Sandpiper)	No tissue collected at this location as birds are not expected to				
B2-Bioaccumulation (Scaup)	forage in this area due to the depth of the water				
C-Chemistry Screening	-	1	1.68	1.68	2

TABLE 6-2

NFD POINT MOLATE MEASUREMENT ENDPOINT RANKS AND WEIGHTS

Attributes	Bulk Sediment Bioassay	Topsmelt SWI Bioassay	Mysid SWI Bioassay	Bio-accumulation	Sediment Chemistry
Degree of Association	**3**	3	3	**2**	**0**
Stressor/Response	**3**	3	3	**1**	**0**
Utility of Measure	4	3	3	**2**	**2**
Quality of Data	5	5	5	5	5
Site Specificity	4	4	4	3	**2**
Sensitivity	**3**	**2**	**2**	3	**0**
Spatial Representativeness	4	4	4	3	**1**
Temporal Representativeness	4	3	4	3	3
Quantitative Measure	5	5	5	**2**	3
Standard Measure	5	**2**	**2**	3	4
Calculated Weight Using MA WOE Scaled Attributes	3.80	3.44	3.48	2.70	1.68

BOLD indicates identified sources of uncertainty (e.g., low ranks for measurement endpoint attributes)

TABLE 6-3

NFD POINT MOLATE AND BIOASSAY WATER QUALITY

		Measured Water Quality Parameter Ranges		
		pH	**Temperature (^{0}C)**	**Salinity ($^0/_{00}$)**
NFD Point Molate Water Quality		Not Collected	14.7 - 20.4	24 - 30
Laboratory	*E. estuarius*	7.7 - 8.7	14.8 - 15.9	19 - 23
Bioassay	*M. bahia*	7.1 - 8.3	24.0 - 27.3	24 - 28
Water Quality	*A. affinis*	7.2 - 8.4	19.5 - 20.9	19 - 23

TABLE 6-4

SUMMARY OF NFD POINT MOLATE ERA CONCLUSIONS

Assessment Endpoint (Section 2.0)	Measurement Endpoint (Section 2.0)	Effects Assessment Results (Section 4.0)	Risk Characterization Results (Section 5.0)	Level of Relative Certainty Based on WOE Weights (Section 7.0)	Primary Sources of Uncertainty (Section 7.0)	Factors Mitigating Uncertainty (Section 7.0)	WOE Results (Section 6.0)	Conclusions
Protection of the benthic invertebrate community associated with NFD Point Molate intertidal	Amphipod Bulk Sediment Bioassay	No effects observed at NFD Point Molate sampling stations	Risk is not indicated	Moderate to High	The use of surrogate species	Surrogate species is well established; well developed San Francisco Bay-specific criteria exist to evaluate results and a second test species (mysid) was tested to evaluate this assessment endpoint.	The WOE Bar Charts show: • All three measurement endpoints have a negative indication of risk at all NFD Point Molate sampling locations except two (T10 and T11). • T10 and T11 have positive findings for the sediment chemistry measurement endpoint; however, sediment chemistry is not an effects-based criterion. • When T10 and T11 are evaluated in the context of the bioasay data, a WOE evaluation indicates that these stations do not indicate risk to the benthic invertebrate community.	• All three measurement endpoints do not indicate risk to the benthic invertebrate community at NFD Point Molate. • Risk from potentially contaminated sediments to the benthic invertebrate community at NFD Point Molate is not indicated. • Uncertainties have been managed and are acceptable and no further data collection is recommended.
Offshore Sediments	Mysid SWI Bioassay	No effects observed at NFD Point Molate sampling stations	Risk is not indicated	Moderate	(1) test species is a surrogate species (2) well developed criteria to evaluate results have not been developed	(1) the use of a well established, sensitive surrogate test species (2) development of a conservative criteria to evaluate results		
	Sediment Chemistry	Criteria not effects-based	Endpoint is not risk-based	Low	(1) criteria is not effects- or risk- based (2) sampling approach is not random	(1) findings used as collaborative information only (2) samples were collected at "worst case" locations		
Protection of the larval fish community associated with offshore eelgrass beds at NFD Point Molate	Topsmelt SWI Bioassay	No effects observed at NFD Point Molate sampling stations	Risk is not indicated	Moderate	(1) performance of the reference location (76% survival) used to evaluate NFD Point Molate results (2) well developed criteria to evaluate results have not been developed	(1) evaluation of NFD Point Molate results using a theoretical reference result (90% survival) (2) development of a conservative criteria to evaluate results.	The WOE Bar Charts show: • Negative indication of risk at all NFD Point Molate sampling locations except two (T10 and T11). • T10 and T11 have positive findings for the sediment chemistry measurement endpoint, however, sediment chemistry is not an effects-based criterion. • When T10 and T11 are evaluated in the context of the bioassay data, a WOE evaluation indicates that these stations do not indicate risk to the larval fish community.	• Both measurement endpoints do not indicate risk to the larval fish community at NFD Point Molate. • Risk from potentially contaminated sediments to the larval fish community at NFD Point Molate is not indicated. • Uncertainties have been managed and are acceptable and no further data collection is recommended.
	Sediment Chemistry	Criteria not effects-based	Endpoint is not risk-based	Low	(1) criteria is not effects- or risk-based (2) sampling approach is not random	(1) findings used as collaborative information only (2) samples were collected at "worst case" locations		
Protection of the shorebird community that forage in the intertidal mud flats at NFD Point Molate	Avian sum PAH dose evaluation	No adverse effects are expected at NFD Point Molate sampling stations	Risk is undetermined	Moderate	(1) inputs to the dose calculation (2) the qualitative criteria used to evaluate doses	Conservative assumptions were used to select inputs to the dose calculation and develop criteria to evaluate doses.	The WOE Bar Charts show for the scaup: • Negative indication of risk at 6 NFD Point Molate sampling locations (T3, T5, T6, T9-1, T9-2 and DL-1). • At two locations (T2 and T11A) the potential for risk is undetermined. At two additional locations (T10 and T11), there are positive findings for the sediment chemistry measurement endpoint.	• Risk to the shorebird community at NFD Point Molate is undetermined.
	Sediment Chemistry	Criteria not effects-based	Endpoint is not risk-based	Low to Moderate	(1) criteria is not effects or risk based (2) sampling approach is not random	(1) findings used as collaborative information only (2) samples were collected at "worst case" locations	The WOE Bar Charts show for the Western sandpiper: • The potential for risk is undetermined at all sampling locations. • At two additional locations (T10 and T11), there are positive findings for the sediment chemistry measurement endpoint.	

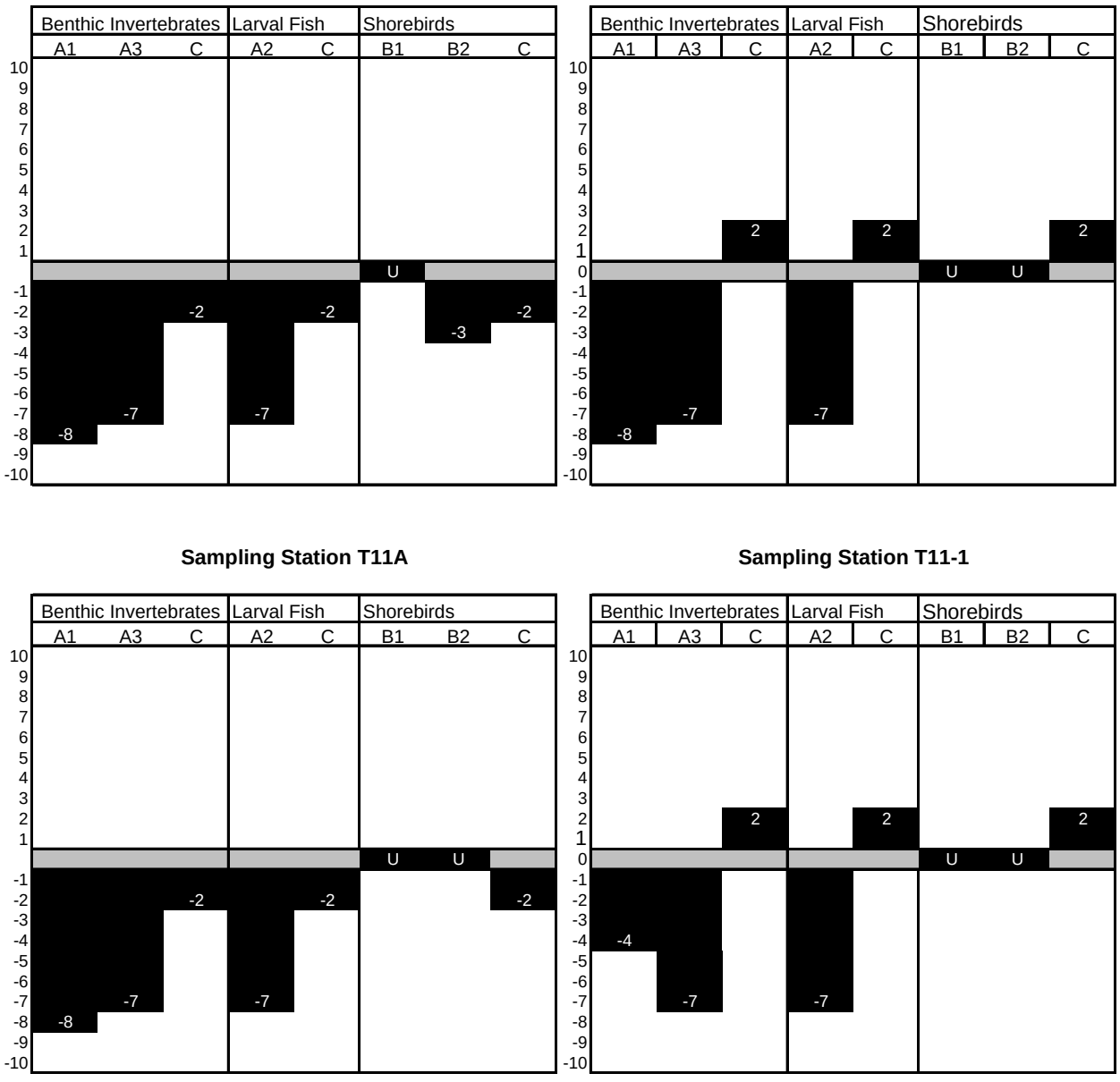


FIGURE 6-1

BAR CHARTS REPRESENTING WOE RANKING SCORES FOR INDIVIDUAL MEASUREMENT ENDPOINTS AT NFD POINT MOLATE SAMPLING LOCATIONS

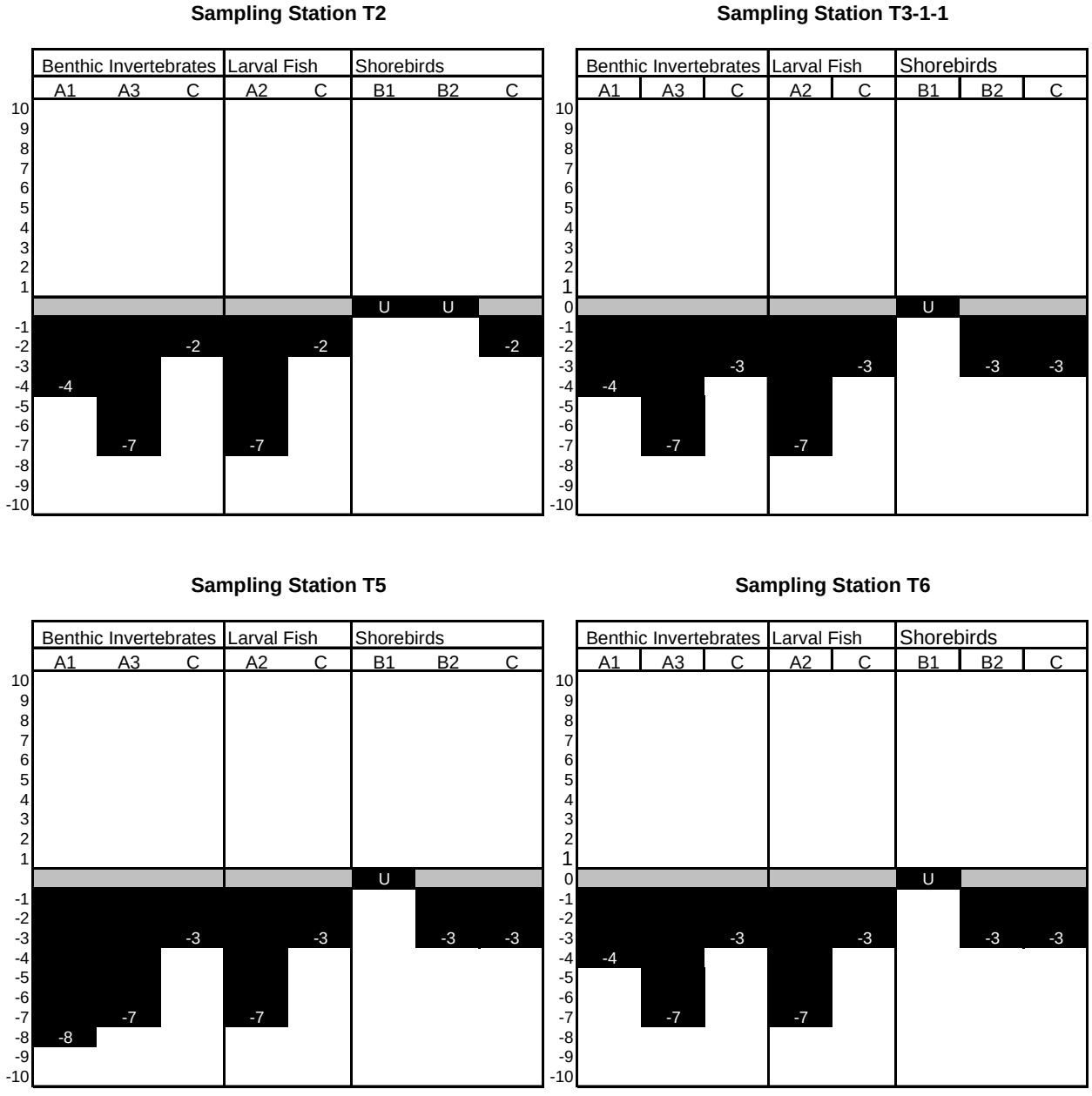
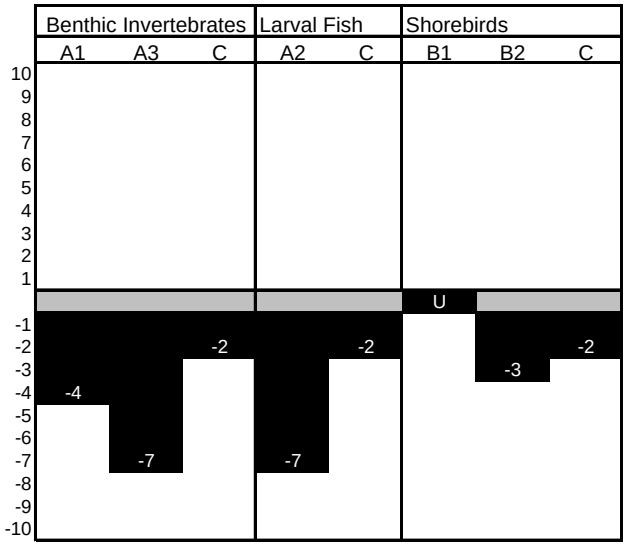


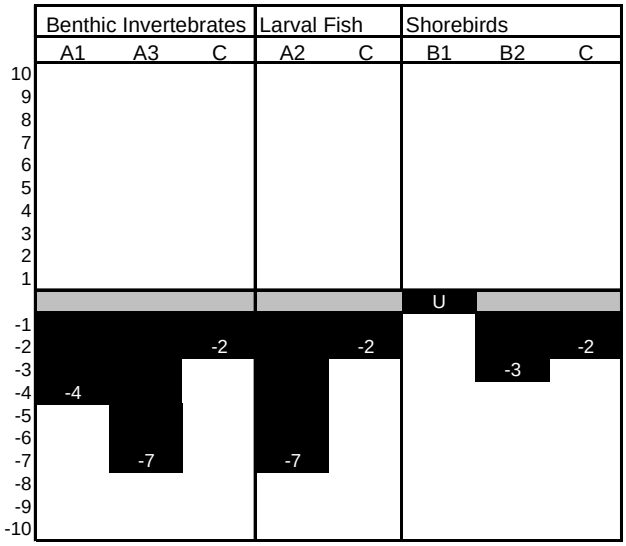
FIGURE 6-1 (Continued)

BAR CHARTS REPRESENTING WOE RANKING SCORES FOR INDIVIDUAL MEASUREMENT ENDPOINTS AT NFD POINT MOLATE SAMPLING LOCATIONS

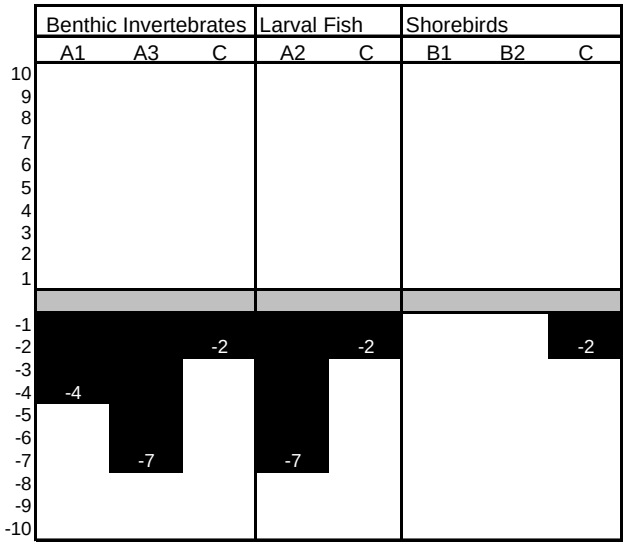
Sampling Station T9-1



Sampling Station T9-2



Sampling Station P1-1



A1=Bulk Sediment Bioassay
A2=SWI (Topsmelt)
A3=SWI (Mysid)
B1-Bioaccumulation (Sandpiper)
B2-Bioaccumulation (Scaup)
C-Chemistry Screening

FIGURE 6-1 (Continued)

BAR CHARTS REPRESENTING WOE RANKING SCORES FOR INDIVIDUAL MEASUREMENT ENDPOINTS AT NFD POINT MOLATE SAMPLING LOCATIONS