Prince William Sound
Regional Citizens' Advisory Council

Contract No. 611.98.1

Long Term Environmental Monitoring Program
Data Analysis of Hydrocarbons in Intertidal Mussels and Marine Sediments, 1993-1996

Submitted by:

J.R. Payne Environmental
317 No. El Camino Real, Suite 201
Encinitas, CA 92024

(760) 942-1015 FAX (760) 942-1036

e-mail: jamesrpayne@compuserve.com

Principal Consultants:

James R. Payne, Ph.D.
William B. Driskell
Dennis C. Lees

Final Report

March 16, 1998

List of Figures

List of Tables

List of Appendices

EXECUTIVE SUMMARY

Prince William Sound RCAC
Long Term Environmental Monitoring Program
1993-1997

Objective

The primary objective of the ongoing Long-Term Environmental Monitoring Program (LTEMP) is to collect data to monitor hydrocarbon pollution due to the oil transportation industry in Prince William Sound and the northern Gulf of Alaska.

Approach

To accomplish this objective, mussel tissues and sediments have been collected for detailed hydrocarbon analyses at nine stations (Figure 1-1, pg 3) from March 1993 through July 1997. Samples include mussel tissues from intertidal habitats along with shallow (5 to 10 meters) and deep (28 to 43 meters) sediments. To support the interpretation of the hydrocarbon data, additional measurements, including lipid content in mussel tissues, shell characteristics, and reproductive state, have been measured on the mussels. The sediment samples have also been characterized for total organic carbon and grain-size distribution.

In the laboratory, the hydrocarbons are chemically extracted from the samples and the polynuclear aromatic hydrocarbons (PAH) are analyzed by selected-ion-monitoring (SIM) gas chromatography/mass spectrometry (GC/MS). Aliphatic hydrocarbons (AHC) are analyzed by flame-ionization-detector gas chromatography (FID-GC). These techniques allow the identification of individual compounds that are characteristic of biogenic hydrocarbons, Alaskan North Slope (ANS) crude oil, *Exxon Valdez* Oil Spill (EVOS) residues, and background signals derived from other sources (e.g., combustion products or the Katalla oil seeps and coal particles transported from outside of the Sound).

The sediment samples were analyzed for both PAH and AHC throughout the program. The mussel tissues were analyzed for both classes of compounds initially, but naturally occurring lipids in the tissues interfered with the aliphatic analyses, so only the PAHs have been analyzed consistently over time. In this executive summary, we will focus primarily on the PAH constituents in several representative samples to show the range of component patterns and concentrations observed. Details of all AHC and PAH analyses are presented on a station-specific basis in Appendix I to this report.

Most crude oils and other sources of PAH that have been introduced to Prince William Sound have specific members of these hydrocarbon groups present in unique ratios

relative to each other. This characteristic allows us to obtain relatively distinctive fingerprint patterns. Figure 3-1 (pg 9) shows the fingerprint (histogram plots) of the observed distribution of the 39 LTEMP target PAH constituents in fresh Alaskan North Slope/EVOS crude oil and three sediment samples that are generally representative of the range of patterns and concentrations observed throughout the program. Figure 3-2 (pg 11) shows the same reference oil contrasted to selected mussel tissue samples extracts. Five main groupings or clusters of peaks are noted in these histograms, including: the naphthalenes, the fluorenes, the phenanthrenes/anthracenes, the dibenzothiophenes, and the chrysenes. Each group includes the parent aromatic component and the C1- through C3- or C4-alkyl-substituted homologues. Each individual bar in the histogram plots represents the absolute concentration of a given component (or group of similar molecular-weight alkyl-substituted components) per gram of oil, sediment, or tissue sample extracted. The histogram plots for the sediment and mussel tissue samples represent the mean of three replicate analyses, and error bars from the standard error of the arithmetic mean associated with each measurement are also shown. The dashed line running across the sediment and tissue sample histogram plots represents the average method detection limit (MDL) for the individual PAH in that sample. It is obviously only a consideration in the lower-concentration samples, such as shallow sediments and mussel tissues from cleaner areas.

For purposes of the LTEMP program, it is useful to distinguish between these five groups of PAH components. The naphthalenes (which are two-ring aromatics) are less persistent in the environment compared to the other groups, and they are subject to weathering from spilled oil by evaporation and dissolution processes. As such, they may or may not be present in the histogram plots of oil-contaminated samples obtained from the environment. The fluorenes, anthracenes, and phenanthrenes (which are all three-ring aromatics) are each more persistent in the environment, and as such, they act as markers to help differentiate among different sources. The dibenzothiophenes (another three-ring compound that also contains sulfur) are important, because they are characteristic of Alaskan North Slope crude oil, but not Cook Inlet or Katalla crude oil. Finally, the four- and five-ring aromatics (including, the chrysenes through benzo(g,h,i)perylene) are important because they can help distinguish between crude oils and refined products (such as diesel oil) that may have been produced from a particular crude oil. They are also representative of combustion by-products. In crude oils, the alkylated PAH components generally predominate over the parent aromatic within a given group, whereas in combustion products, the parent aromatic compound is more predominant.

The relative PAH patterns presented in Figures 3-1 and 3-2 will be considered briefly below; detailed comparisons of different sources, sediment depths, grainsize dependence, etc. are presented on a site-specific basis in Appendix I. For the purpose of this discussion, it is sufficient to recognize that these different constituents weather (evaporate and dissolve) to different degrees after oil is released to the environment. As such, knowledge of how these patterns change with time and exposure conditions is also important in identifying contaminants in the different samples examined.

Results

Overall concentrations and general sources

In the 4.5 years of data collection for this program, there have been some changes in the sampling design regarding station coverage. As a result, sediment data are not always available for all stations at both depths or seasons. Mussel tissue collections, however, have been essentially continuous for the March and July samplings at all stations for the entire period. Between the sediment and mussel tissue collections, sample coverage is sufficient to allow time trends and sources of individual hydrocarbons at the different stations to be identified.

All hydrocarbon concentrations in both sediments and tissues are generally very low. PAH concentrations in the cleaner sites for deep sediments range from less than 30 nanograms per gram (ng/g) dry weight of sediment at Aialik Bay to over 500 ng/g dry weight in Sleepy Bay. Shallow sediments have generally even lower PAH levels, ranging from less than 10 ng/g dry weight at Knowles Head to approximately 400 ng/g dry weight at Sleepy Bay. Average PAH hydrocarbon burdens in the mussel tissues ranged from less than 130 ng/g dry weight at Aialik Bay to over 510 ng/g dry weight at Alyeska Marine Terminal in the Port of Valdez.

The PAH patterns in the sediments and tissues examined in this program reflect several "background" sources, including the Katalla oil seeps and/or coal particles derived from the rivers east of Prince William Sound, as well as oil-transportation activities associated with the Alyeska Marine Terminal in the Port of Valdez. At this time, there is a debate in the scientific literature as to whether the natural "background" hydrocarbons are actually derived from oil seeps or from coal particles. However, for the purposes of this program, it is sufficient that these fingerprints can be distinguished from the pattern generated from Alaskan North Slope crude oil introduced from present-day activities or weathered EVOS residues that are still present at a few locations.

The histogram plots for the sediment samples presented in Figure 3-1 illustrate several interesting features. The individual PAH concentrations for the Disk Island deep sediments are very low (generally 3-12 ng/g dry weight), and the precision as reflected by the small standard error bars, is very good. Several of the components are below the average individual component method detection limit (MDL) of 0.7 ng/g dry weight, but the precision is again very tight, and a characteristic pattern classified as being from sources such as coal or seep oil from outside Prince William Sound is obtained (Short and Babcock, 1996). The individual PAH concentrations for the Disk Island shallow sediments are approximately an order of magnitude lower than the deep sediments, and most of them are, in fact, below the average individual MDL. Nevertheless, the precision among the triplicate measurements is very good, and it is possible to identify the same general background pattern as noted for the deeper sediments. The Disk Island intertidal sediments collected during cruise 6 (July 1995) show the classic weathered PAH pattern

associated with EVOS residues. It appears that essentially 100% of these PAHs are derived from very well weathered *Exxon Valdez* oil residues (Page et al., 1995, 1996). Discussions of these and other features are presented in Appendix I.

The histogram plots obtained on the representative mussel tissue extracts shown in Figure 3-2 also tell an interesting story. In both the Sheep Bay cruise 3 (March 94) and the Gold Creek cruise 6 (July 96) samples, most of the individual analytes are below the average individual component MDL of 12 ng/g dry weight. Nevertheless, the precision obtained on the triplicate samples is very good, and the same general pattern is obtained in both samples. In fact, this pattern is characteristic of most of the mussel samples obtained from cleaner areas throughout the study area where few, if any alkylated PAH derived from the more common oil sources are observed. Many of the constituents have been identified by NOAA (1997) as combustion-derived PAH by-products from burning oil. At the same time, Bence and Burns (1995) describe a similar profile as "procedural artifacts." In actual fact, many of these components do routinely show up in this exact pattern in procedural and field blanks analyzed during this program. They are, however, generally present at even lower concentrations.

The point to be emphasized here is that the levels shown in the top two histograms in Figure 3-2 are extremely low, and that when a pulse of oil is released, it is easily detected, as shown by the histogram obtained from the Gold Creek mussels collected during cruise 9 (March 97). In this instance, the characteristic pattern of relatively fresh ANS oil can be observed and potentially traced back to the Alyeska Ballast Water Treatment Plant spill that occurred in January 1997. The relative contribution of PAHs from ANS crude oil to the overall PAH burden may be as high as 50%.

Analysis of geographical and time-series trends

Because there is natural variability among samples from a site, the LTEMP program collected three replicate samples of sediment and tissue from each site. The results from each type of sample were averaged to generate a total aliphatic hydrocarbon (TAHC) and/or total aromatic hydrocarbon (TPAH) concentration for each station, at each time. In addition, as described in the previous section, histogram plots for the individual aliphatic and aromatic compounds were generated for all the samples for the project chemist to evaluate. Although there is generally a fair amount of natural variability among samples within any given station at any given time, the data generated in this program are quite precise, and they allow evaluations of geographic and time-series trends among the stations or over time at a single station.

To aid in analyzing all of the available data from this program, a new empirical value, we named the "CRUDE" index, was developed to combine into a single value several individual parameters and ratios used to identify petroleum patterns within the data. The CRUDE index functions as a summation of TPAH, TAHC, and the unresolved complex mixture (UCM) of the aliphatic hydrocarbon analyses with each term weighted to emphasize the petrogenic fraction. The analyst (and reader) can then deal with a single variate versus a multivariate analysis.

Figure 4-1 (pg 28) presents the CRUDE index values obtained from the sediment samples collected in the control sites at Aialik Bay, Gold Creek, and Sheep Bay; the EVOS-impacted sites at Disk Island, Shuyak Harbor, Sleepy Bay, and Windy Bay; and sites associated with tanker activities, Alyeska Marine Terminal and Knowles Head anchorage. Standard error bars reflecting the scatter associated with each triplicate measurement are also printed on top of each sample presented in the figure. As noted earlier, sediment samples were collected at deep and shallow stations. Therefore, in the figure, station identifications are denoted as DII-M-2 or DII-S-3, etc. DII-M-2 stands for DIsk Island, Mid-depth sediment, cruise 2, and DII-S-3 represents DIsk Island, deep Sediment, cruise 3, etc.

As shown by Figure 4-1, relatively flat and extremely low-level CRUDE index values are obtained for the deep sediments at Aialik Bay and Gold Creek (control stations); the mid-depth sediments within Windy Bay and Disk Island (EVOS-impacted stations); and finally in the mid-depth sediments at Knowles Head (tanker-affected area). At these stations, there was very little change observed in the absolute hydrocarbon concentrations and little apparent change in the patterns associated with the histogram plots generated for each station over time. Likewise, these stations exhibited little or no evidence of ANS or EVOS derived oil, and only extremely low-level background hydrocarbons from the petrogenic or coal patterns were noted.

Some changes in sediment hydrocarbon burdens were suggested over time by the increases in the CRUDE index values at Sheep Bay (mid-depth and deep), at Disk Island (deep), and at Sleepy Bay (mid-depth and deep). Likewise, very high variability and much higher absolute concentrations of petroleum-derived hydrocarbons were noted in the deep sediments at the Alyeska Marine Terminal.

Figure 4-4 (pg 38) presents the data generated for the *Mytilus* Petrogenic index, which is generated from the sums of individual compounds that are particularly characteristic of PAHs derived from petroleum as opposed to combustion sources. As with the CRUDE index plot discussed above, standard error bars reflecting the scatter associated with each triplicate measurement appear on top of each compound in the sample. In this case, there are time-series changes and patterns noted for the mussel samples collected at essentially every station. The relative magnitude of the error bars associated with each triplicate measurement is very small, however, when compared to the overall change in *Mytilus* Petrogenic index values over time. Therefore, the observed trends are believed to reflect real changes in the field, and not artifacts of the analytical method or collection procedure. The patterns observed at several of these stations can, in fact, be correlated with spill events or clean-up activities that have occurred in Port Valdez or Prince William Sound since 1993 (see below).

Hot spots and areas of high variability

Table 5-1 (pg 83) presents the major observations from the LTEMP program. This table was generated after detailed examination of every aliphatic and aromatic hydrocarbon histogram plot for every sample, the trends from the CRUDE and *Mytilus* Petrogenic

indices, and the fingerprint ratios that are characteristic of different oil sources. From the data summary, highly variable stations or so-called "hot spots", indicating higher oil concentrations, were noted. A subset of those hot spots was then examined for evidence of ANS or EVOS oil, and identified in the table. From these analyses, the following distribution of ANS or EVOS-related oil were observed:

- Alaska Marine Terminal – ANS oil was detected in deep sediments for all nine cruises; mussels showed evidence of ANS oil for cruises 1, 3, 4, 5, 7, 8, and 9.
- Disk Island – No ANS or EVOS oil was detected to any significant extent in the deep or shallow sediments; mussels exhibited evidence of EVOS oil in cruises 3, 4, 5, 8, and 9; intertidal sediments (opportunistically collected when weathered oil was observed on the beach during mussel collection) showed significant quantities of EVOS oil in cruises 6 and 8.
- Gold Creek – ANS oil was observed only once in deep sediments during cruise 4; mussels showed ANS oil in cruises 1, 3, 5, and 9.
- Knowles Head -- no evidence of Alaskan North Slope crude was noted in either the anchorage or shallow sediment locations; however, ANS crude was detected in the mussel samples collected during cruise 8.
- Sheep Bay -- no samples showed any evidence of ANS or EVOS oil for either sediments or mussel tissue.
- Shuyak Harbor – ANS or EVOS oil was noted in the deep sediment for cruise 3 only; mussel tissue showed no contamination from ANS or EVOS oil.
- Sleepy Bay -- there was no evidence of ANS or EVOS oil in the deep sediment samples; however, the shallow sediments showed positive hits for ANS or EVOS-derived oil during cruises 4, 6, and 7. Mussel samples showed evidence of ANS or EVOS oil in cruises 1, 3, 4, 5, and 9.
- Windy Bay – ANS or EVOS-derived oil was observed in the deep sediments during cruises 3, 4, 6, and 8, but not in any of the shallow sediments. ANS or EVOS oil was detected in the mussel samples only during cruise 9.

Table 5-1 also lists the range of values obtained for the relative percent ANS or EVOS contribution to the total aromatic hydrocarbon burdens in the different sample matrices at the different stations. The relative percent contributions range from nondetect to 100 percent of the total PAH measured. It should be remembered, however, that the total PAH levels in most of these samples were extremely low. Therefore, although the relative percent ANS or EVOS oil for any given station may have been high, the absolute value for the concentration of residual oil itself was extremely low.

Correlation of mussel hydrocarbon values with known events

On initial examination, the mussel hydrocarbon patterns may appear wildly variable with no apparent trend or explanation. However, the trends observed in Figure 4-4 can be correlated with a chronology of documented events that have occurred within Port Valdez and Prince William Sound since 1993. In May of 1994, the *Eastern Lion* oil spill occurred at the Alyeska Marine Terminal during loading operations. Mussel samples collected at the time of the spill showed extremely high levels of hydrocarbons at the Alyeska Marine Terminal station, which is located near berth 5, the site of the spill. Elevated levels were still noted in the mussel tissues at Alyeska Marine Terminal during

cruise 4 (July 1994), and a strong signal was observed in the Gold Creek samples at the same time.

A similar spike in the mussel contamination from a weathered ANS source was also noted in the samples collected at Disk Island during cruise 4 (July of 1994), and at first it might seem plausible to speculate that it too might be from the *Eastern Lion* oil spill. However, sheens released from mussel bed cleaning operations at Disk Island just prior to the RCAC samplings are a much more likely source.

After the July 1994 events, hydrocarbon levels in the mussels at all stations dropped to uniformly low values by cruise 6 (July 1995). The PAH histogram pattern for the mussels collected from Gold Creek in July 1995 (Figure 3-2) is indicative of the extremely low background signal observed in mussel samples throughout Prince William Sound at that time. As noted earlier, this pattern is identified as either being characteristic of the by-products associated with the combustion of oil or as a low-level procedural artifact of the sampling and measurement program. It also shows up consistently at low-level sites in other monitoring efforts, such as the NOAA Status and Trends program.

Examination of the *Mytilus* Petrogenic index plot shows another increase at all stations during cruises 8 and 9 (July 1996 and March 1997). The profiles obtained in these samples are again consistent with those observed for Alaskan North Slope crude oil at Alyeska Marine Terminal, Gold Creek, and Disk Island. The cruise 9 profiles for the increase observed at Sleepy Bay and Knowles Head, however, are not consistent with the source being Alaskan North Slope crude oil. One possible source for a newly arising signal observed in the last sampling interval (at least for Alyeska Marine Terminal and Gold Creek), would be from the Alyeska Ballast Water Treatment Plant (BWTP) spill, which occurred in January of 1997.

The histogram profiles associated with the increases in hydrocarbon concentrations at Sleepy Bay during cruises 3, 4, and 5 show that the source is consistent with EVOS or more recent releases of Alaskan North Slope crude oil.

The interpretation of the *Mytilus* Petrogenic index pattern at Windy Bay is somewhat more complicated. A mixed source is indicated, including contributions from aromatics that look like they could be derived from Bunker C or No. 6 fuel oil. In addition, the contributions from biogenic hydrocarbons (plant waxes and natural oils), as measured in the sediments, are higher at Windy Bay than at any other site. This site also contains traces of Alaskan North Slope or EVOS-related oil that were detected in the deep sediments during cruises 3, 4, 6, and 8. However, hydrocarbons associated with ANS or EVOS oil were not observed in the shallow sediments. ANS or EVOS oil was only observed in the mussel samples collected during cruise 9 at Windy Bay. Kinnetics personnel (the field samplers) suggest that logging operations in the area may be a new possible source of both biogenic and petrogenic hydrocarbons.

Conclusions and Recommendations

In sections 5 and 6 of this report, we present further conclusions, assessments of the program's effectiveness, and recommended adjustments. Suggested changes include comments on:

- adjusting the sampling plan to include more sites,
- modifying the statistical criteria,
- adding intertidal sediment samples,
- rectifying MDL problems in the laboratory analyses,
- paying closer attention to field and procedural blank contamination problems,
- reinstating aliphatic hydrocarbon analyses in mussel tissue samples,
- tightening field sample procedures regarding sampling depth and mussel size,
- dropping mussel lipid corrections, seasonal sampling, and unnecessary shell measurements, and
- sampling and analyzing potential background sources with common laboratory methods.

Final Report

Analysis of Hydrocarbon Data from the Long-Term Environmental Monitoring Program in Intertidal Mussels and Marine Sediments, 1993-1996

1. Introduction

1.1 Scope of the Report

This report is not intended to be a comprehensive treatise on the overall state of hydrocarbon contamination within Prince William Sound (PWS). We have limited our focus (by design and contract requirements/directives from the PWS RCAC) to identification of recent introductions of Alaskan North Slope (ANS) crude oil or *Exxon Valdez* Oil Spill (EVOS) residues at the specific sites monitored in the program. We do not attempt to differentiate among the other potential sources of background hydrocarbons which are found throughout Prince William Sound (Short and Babcock, 1995; Page et al., 1995, 1996; and Bence and Burns, 1995, among others) although we do note their presence.

1.2 Report Organization and Contents

In the report that follows, the objectives and designs of the LTEMP program are first briefly considered, and then methods are presented for data analysis and overall program evaluation. Because most of these methods or approaches have been used before, extensive details on their development and implementation are not presented unless we've modified their formulation. In the results section, temporal and geographic variations in hydrocarbons are discussed along with evaluation of potential sources. We believe there is good evidence that the program has successfully tracked at least three significant and several minor hydrocarbon releases within Port Valdez and in Prince William Sound. In this regard, the program appears to be achieving its goal of detecting low-level changes in hydrocarbon burdens and in identifying the presence or absence of Alaskan North Slope crude in sentinel organisms.

General discussions of overall trends in sediments and mussel tissues are presented along with statistical considerations, and then field sampling and chemistry issues are briefly considered.

The conclusions and integration section consolidates our findings and presents several hypotheses for the observed distributions and patterns measured over the course of the program. Finally, recommendations for redirection and/or modification of the LTEMP are broken out for consideration.

This is a summary stand-alone document that is intended to present a concise technical overview of the data obtained from the LTEMP. The report includes site-specific summary narratives integrating the results of detailed analyses of <u>all</u> hydrocarbon data (PAH and AHC histogram plots for every sample), sediment grain size/depth

considerations, method detection limits, and field/procedural blanks, etc., for all stations and all sampling times. To enhance the readability of the report, the comprehensive summaries are presented by individual station in Appendix I.

1.3 Discussion of Long-Term Environmental Monitoring Project

The primary objective of the ongoing Long-term Environmental Monitoring Program (LTEMP) is to collect "standardized measurements of hydrocarbon background in the EVOS region as long as oil flows through the pipeline". The measured variables include polycyclic (or polynuclear) aromatic and aliphatic hydrocarbon levels (PAH and AHC) in mussel tissues and sediments, sediment grain size and organic carbon content, and several measurements of size and condition from individual mussels (*Mytilus trossulus*) and pooled samples.

1.3.1 Analyses of Aliphatic and Aromatic Hydrocarbons in Sediments

PAH and AHC levels including UCM (unresolved complex mixture) were measured for all sediment samples by the Geochemical and Environmental Research Group (GERG) oil chemistry laboratory at Texas A&M University. Sampling and analytical methods are patterned after the protocols developed by the National Oceanic and Atmospheric Administration (NOAA) Status and Trends (Mussel Watch) Program (NOAA 1993), and they are fully detailed in the annual Monitoring Reports prepared by Kinnetic Laboratories, Inc. (KLI) and GERG. This study examines the results of roughly 300 field samples collected from within Prince William Sound and the surrounding region (Figure 1-1) in addition to the laboratory quality control results. The frequency of site visits is summarized in Table 1-1.

1.3.2 Aromatic Hydrocarbon Analyses in Mussel Tissues

Initially, both types of hydrocarbons (PAHs and AHCs) were measured in mussel tissues. PAHs were consistently measured in all sampling events at all nine stations except for one, the initial cruise 1 at Shuyak Harbor (Table 1-1). AHCs were also measured in tissue samples from all stations in 1993 and 1994, but were discontinued in subsequent years because of interference from lipids in the tissues. Because the aliphatic data set for tissues is incomplete, it is not considered in the analyses of mussel data presented in this report. The data set for PAHs is used for evaluation of seasonal and long-term patterns and determination of stations with high levels and high variability in hydrocarbon concentrations.

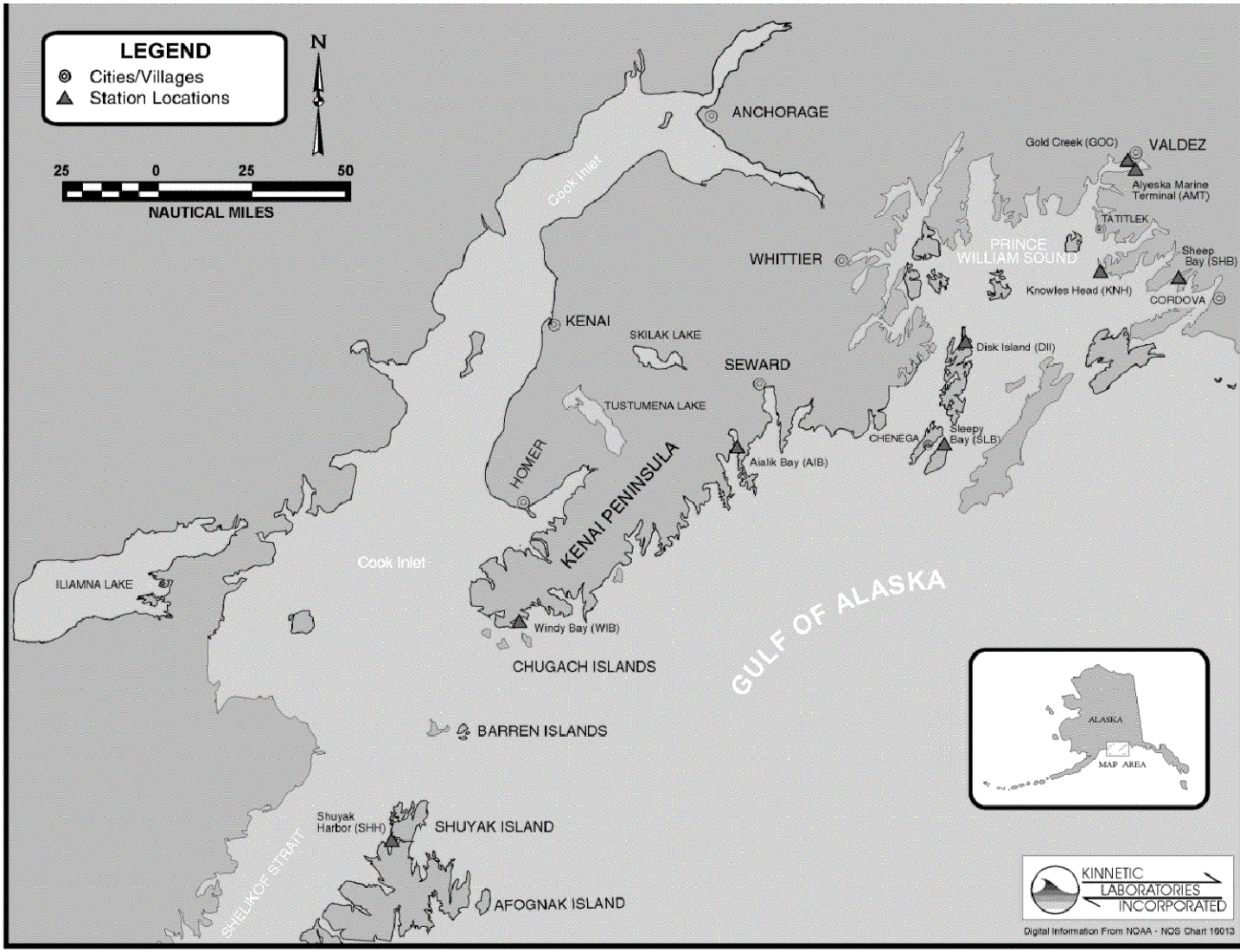


Figure 1. LTEMP Station Locations (Overall Study Area).

Table 1-1 Sampling plan overview, visits by depth and cruise.

		Cruise No.									
		1	2	3	4	5	6	7	8	9	Total Site
Station	Depth	Mar-93	Jul-93	Mar-94	Jul-94	Mar-95	Jul-95	Mar-96	Jul-96	Mar-97	Visits
AIB	B	x	x	x	x	x	x	x	x	x	9
	S	x	x	x	x		x		x		6
AMT	B	x	x	x	x	x	x	x	x	x	9
	S	x	x	x	x	x	x	x	x	x	9
DII	B	x	x	x	x	x	x	x	x	x	9
	I						x		x		2
	M				x	x	x	x	x	x	6
	S	x	x	x	x		x		x		6
GOC	B	x	x	x	x	x	x	x	x	x	9
	S	x	x	x	x	x	x	x	x	x	9
KNH	B	x	x	x	x	x	x	x	x	x	9
	A		x	x	x		x		x		5
	M					x	x	x	x	x	5
PEP	B	x									1
SHB	B	x	x	x	x	x	x	x	x	x	9
	M				x	x	x	x	x	x	6
	S	x	x	x	x		x		x		6
SHH	B		x	x	x	x	x	x	x	x	8
	M					x	x	x	x	x	5
	S		x	x	x		x		x		5
SLB	B	x	x	x	x	x	x	x	x	x	9
	M				x	x	x	x	x	x	6
	S	x	x	x	x		x		x		6
WIB	B	x	x	x	x	x	x	x	x	x	9
	M					x	x	x	x	x	5
	S	x	x	x	x		x		x		6
Cruise Totals		16	18	18	21	17	25	17	25	17	174

Depth Keys
M = shallow (mid) depth sediments
S = deep depth sediments
B = intertidal mussel tissues
A = tanker anchorage sediments
I = intertidal sediment core

Lipid concentrations have been measured in the tissues of all mussels analyzed for PAH to date, and corrections based on lipid content are considered in this report. This measurement is sometimes helpful in normalizing hydrocarbon concentrations in tissues of some organisms because of the lipophilic nature of hydrocarbons (i.e., PAH levels may be strongly correlated with the individual's fat stores). The efficacy of this approach is considered in this report.

1.3.3 Size, Gonadal, and Non-Gonadal Measurements

Several measurements were taken from individual mussels to assess their size and condition. Shell length and volume were measured to provide an indication of the size of the mussels included in the samples. Weight and volume of gonadal tissue were measured to provide an indication of reproductive condition. Weight and volume of non-gonadal (somatic) tissues were also measured separately. Total tissue weights for the

mussels were determined by combining the weights for gonadal and somatic tissues, presumably to provide an indication of the general condition when standardized by shell length or volume.

2. Objectives

2.1 Program Goals

In view of the RCAC program's primary interests regarding hydrocarbon contamination, this report evaluates the suitability of the measured variables for establishing baseline conditions and their sensitivity in detecting new or historic events. By RCAC guidance, evaluation of the biological measurements is of lesser concern than evaluation of the chemical measurements in this project.

2.2 Objectives

2.2.1. Primary Objectives

- Analyze the 1993-97 LTEMP data for temporal (seasonal and annual) and geographical (among sampling stations) variation in hydrocarbon concentrations and signatures, comparing those signatures to known source signatures.

- Identify sampling locations where concentrations of hydrocarbons in mussel tissues and sediments are high or highly variable.

- Evaluate LTEMP methodologies and compare them with those used in other mussel and sediment hydrocarbon studies in Prince William Sound and the Gulf of Alaska.

- Produce a report documenting the findings of the program.

- Assist RCAC's Scientific Advisory Committee (SAC) and LTEMP Project Manager in Developing Recommendations for the LTEMP Sampling Plan for Future Years.

2.2.2 Secondary Objective

Identify, Define, and Describe Environmental Variables Potentially Affecting Hydrocarbon Concentrations.

3. Methods Used for LTEMP Data and Overall Program Evaluation

3.1 Hydrocarbon Characterization and Analysis

3.1.1 Petroleum Hydrocarbon Parameterization and Representative Ratios (portions adapted from KLI, 1997)

Petroleum contains monoaromatic and polycyclic aromatic hydrocarbons (PAH), both of which can be toxic to organisms. Monoaromatic hydrocarbons such as benzene, toluene, and xylene(s) are highly volatile and are quickly lost into the environment through evaporation and dissolution processes (Payne et al., 1983, 1984, 1991a,b,c; Payne and McNabb, Jr., 1984). These compounds do not persist in the marine environment for long periods of time and have not been measured in this study. Petroleum also contains an extensive suite of PAHs, and the amount and composition of the PAH fraction can be effectively used as a tracer of petroleum contamination. In general, PAHs are more resistant to microbial breakdown than many aliphatic hydrocarbons, and thus, they tend to persist in the environment longer. PAHs are also toxic and serve as an indication of exposure in organisms. Based on consideration of the petroleum chemistry, biological hydrocarbons (i.e., analytic interferences), and toxicological effects, aliphatic hydrocarbons (AHC) and PAH were chosen as the preferred organic tracers of petroleum contamination in PWS (KLI 1997).

Polycyclic aromatic hydrocarbons are generally divided into three main sources: biogenic, petrogenic, and pyrogenic. The most abundant biogenic PAH is perylene, which is believed to be formed during the early stages of diagenesis -- the bacteriological breakdown of organic matter in marine sediments (Venkatesan, 1988). High concentrations of perylene are often observed in high depositional areas receiving significant concentrations of organic material. Perylene can be easily differentiated from the PAHs in petroleum, and because of its biogenic origin, it is usually not classified as a petrogenic PAH. As a result, it has been excluded from the summation of TPAH in the LTEMP data analysis.

Petrogenic PAHs include all of those commonly identified in crude oil and its refined products. While hydrocarbons are ubiquitous in the marine environment, petrogenic hydrocarbons can be individually recognized and are often used as tracers of oil contamination (Brassell et al., 1978; Boehm and Requejo, 1988; Kennicutt and Comet, 1992). Potential sources of petrogenic PAHs in the LTEMP study area include: Alaskan North Slope (ANS) crude including EVOS oil residues; Cook Inlet crude; oil products from the Alyeska Marine Terminal (not necessarily ANS); Katalla, Yakataga, and other eastern Gulf of Alaska seep oils; and refined petroleum products that have made their way into the marine environment. ANS crude consists of a mixture of petroleum from the production fields on the Alaskan North Slope, including Prudhoe Bay, Kuparuk, Endicott, and Lisburne, and it exhibits a fingerprint that is quite distinct from that of oil found in other geographic areas. The EVOS of March 1989 consisted of ANS crude, which over time has weathered to produce a significantly different fingerprint than that of

fresh ANS crude. Petroleum that originates from natural seeps in the Gulf of Alaska contributes to the natural (or “background") hydrocarbons in the study area, and these also exhibit a distinctly different fingerprint.

Other petroleum products that may have been introduced into the marine environment in PWS include oil products from source locations other than Alaska. For example, the Great Alaskan Earthquake of 1964 and the resultant tsunamis caused the introduction of fuel oil and asphalt made from California source oils into Port Valdez, and subsequently into PWS (Kvenvolden et al., 1995). These authors noted that residues of these California-sourced products have been found throughout the northern and western parts of PWS, typically in the form of tar balls found on beaches at the high tide line.

For purposes of the LTEMP program, it is useful to distinguish between five main groups of PAH components. The naphthalenes (which are two-ring aromatics) are less persistent in the environment compared to the other higher-molecular-weight groups, and they are subject to loss from spilled oil by evaporation and dissolution weathering. As such, they may or may not be present in the histogram plots of oil-contaminated samples obtained from the environment. The fluorenes, anthracenes, and phenanthrenes (which are all three-ring aromatics) are each more persistent in the environment, and as such, they can act as markers to help differentiate among different sources. The dibenzothiophenes (another three-ring compound that also contains sulfur) are important, because they are characteristic of Alaskan North Slope crude oil, but not Cook Inlet or Katalla crude oil. Finally, the four- and five-ring aromatics (including, the chrysenes through benzo(g,h,i)perylene) are important because: 1) they can help distinguish between crude oils and refined products (such as diesel oil) that may have been produced from a particular crude oil; and 2) they are also representative of combustion by-products.

Figure 3-1 presents example histogram plots of the observed distribution of the 39 target LTEMP PAH constituents in fresh Alaskan North Slope/EVOS crude oil and three sediment samples that are generally representative of the range of patterns and concentrations observed throughout the program. The histograms for the Disk Island deep and shallow subtidal sediments are representative of the pattern classified as being from sources such as coal or seep oil from outside Prince William Sound, and the histogram for the Disk Island intertidal sediment is representative of the classic weathered PAH pattern associated with EVOS residues. In comparison, the fresh ANS/EVOS oil standard shows the alkylated naphthalenes as the most prominent constituents, with lower relative concentrations of the other alkylated PAH. When examining the histograms from individual stations, it is important to note that petrogenic PAHs have a characteristic fingerprint where the parent compounds (i.e., C_0-naphthalene, C_0-fluorene, C_0-phenanthrene, C_0-dibenzothiophene, etc.) are usually at lower concentrations than their alkyl-substituted homologues. With evaporation/dissolution weathering, these lower-molecular-weight components are further removed, generating a characteristic “water-washed profile” where the C0<C1<C2<C3 within each PAH group. Eventually, with continued weathering, only the alkylated phenanthrenes/anthracenes, dibenzothiophenes, and chrysenes persist at very characteristic and source-specific ratios in the remaining oil residues.

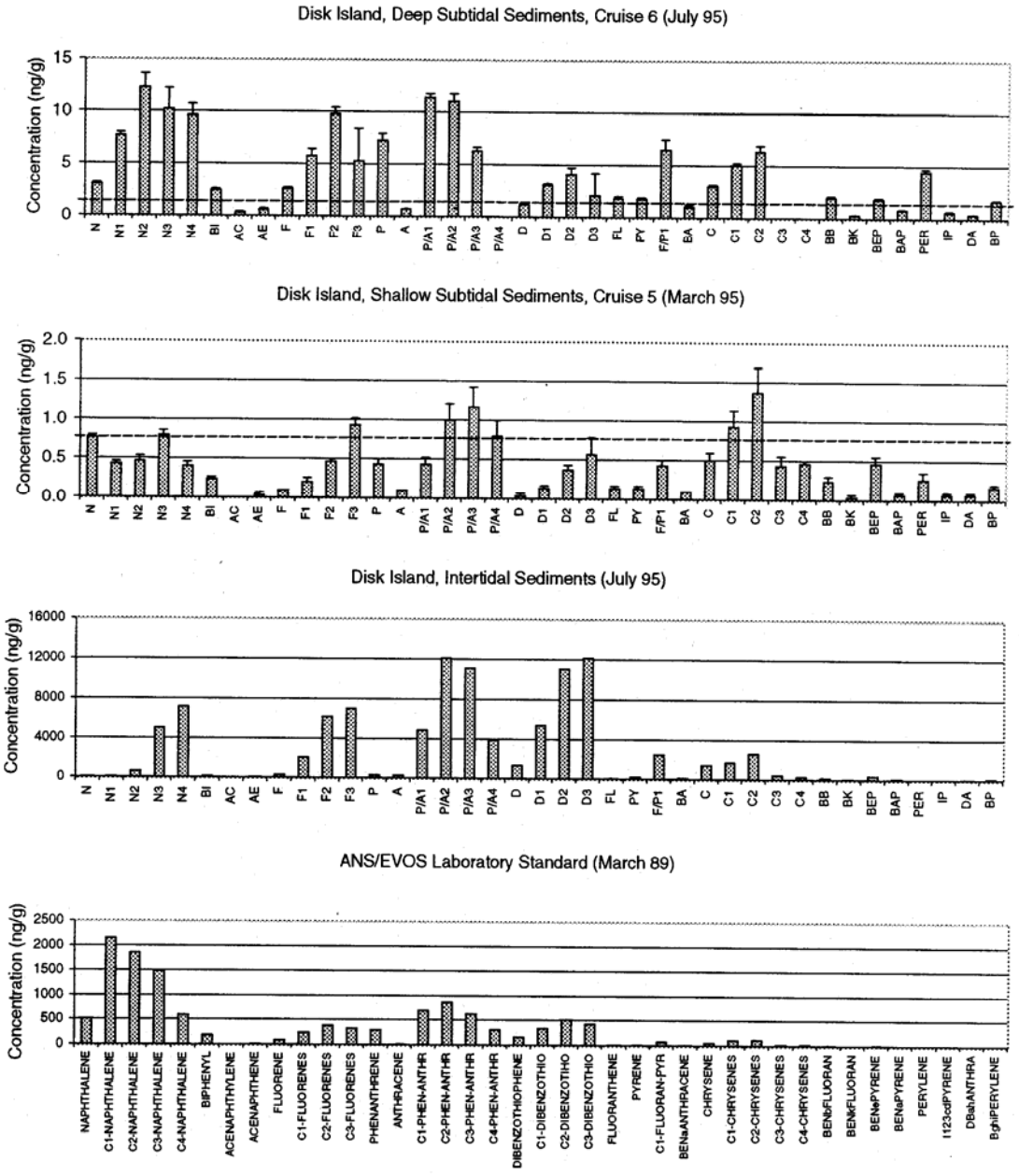


Figure 3-1. Histogram plots showing the observed distribution of 39 target PAH constituents in representative sediment samples and fresh ANS/EVOS oil. Concentrations are in ng/g dry weight of sediment or oil extracted (note the different scales). The histograms for the deep and shallow sediment samples represent the mean of three replicate analyses, and the "error bars" shown above each component represent the standard error of the arithmetic mean for each measurement. The dashed horizontal line in the plots represents the average method detection limit (MDL) for the individual PAH in a sample.

Figure 3-2 shows the histogram from the same reference ANS/EVOS oil contrasted to examples obtained from selected mussel tissue samples that are representative of the concentrations and patterns observed during the program. In both the Sheep Bay cruise 3 (March 94) and the Gold Creek cruise 6 (July 96) tissue samples, most of the analytes are below the average individual component method detection limit (MDL). This pattern was often observed in the mussel samples obtained from cleaner areas throughout the study area, where few alkyl-substituted PAH derived from more common oil sources were observed. At these low levels, however, it was easy to detect a signal when a pulse of oil was released, and this is shown by the histogram obtained from the Gold Creek mussels collected during cruise 9 (March 97). In this instance the characteristic pattern of relatively fresh ANS oil can be readily observed and potentially traced back to the Alyeska Ballast Water Treatment Plant spill that occurred in January 1997. Additional discussions of these figures and similar profiles for other locations are presented on a site-specific basis in Appendix I.

Pyrogenic PAHs come from combustion sources including atmospheric fallout and surface runoff from the burning of fossil fuels (diesel, heating oil, gasoline, etc.) and from other pyrogenic sources such as forest fires and camp fires. Creosote, which is used to preserve wood pilings, is also usually included in this category. Pyrogenic PAHs are characterized by high molecular weight PAHs, greater than C3-dibenzothiophene, and by high concentrations of the parent compounds compared to their alkyl homologues. A typical pattern for pyrogenic PAHs is decreasing concentration with increasing molecular weight within a group, i.e., C0>C1>C2>C3>C4. It has been noted, however, that the PAH in diesel soot has primarily a petrogenic signature (Bence and Burns, 1995).

In contrast to the PAHs, aliphatic and aromatic hydrocarbons can account for more than 70 percent of petroleum by weight. Aliphatic hydrocarbons can also be synthesized by organisms (both planktonic and terrestrial). Crude petroleum contains an homologous series of n-alkanes with one to more than 30 carbons with odd and even n-alkanes present in nearly equal amounts, whereas biogenic hydrocarbons produced by living organisms preferentially contain specific suites of normal alkanes with odd numbers of carbons from 15 to 33. Petroleum also contains a complex mixture of branched and cyclic compounds generally not found in organisms, although the latter may be found as degradation products in some bacteria. This complex mixture can include oxygenated compounds that produce an unresolved complex mixture of compounds (the UCM) on the gas chromatograph when petroleum is extensively biodegraded. The presence and amount of the UCM can be a diagnostic indicator of heavily weathered petroleum contamination.

Through the investigation of petroleum weathering and persistence in the marine environment, several investigators have developed an innovative suite of characteristic ratios, sums, and other indices to aid them in identifying petroleum signatures and distinguishing petrogenic from pyrogenic and biogenic sources. Table 3-1 summarizes a number of the different factors that have been used in the LTEMP data analysis.

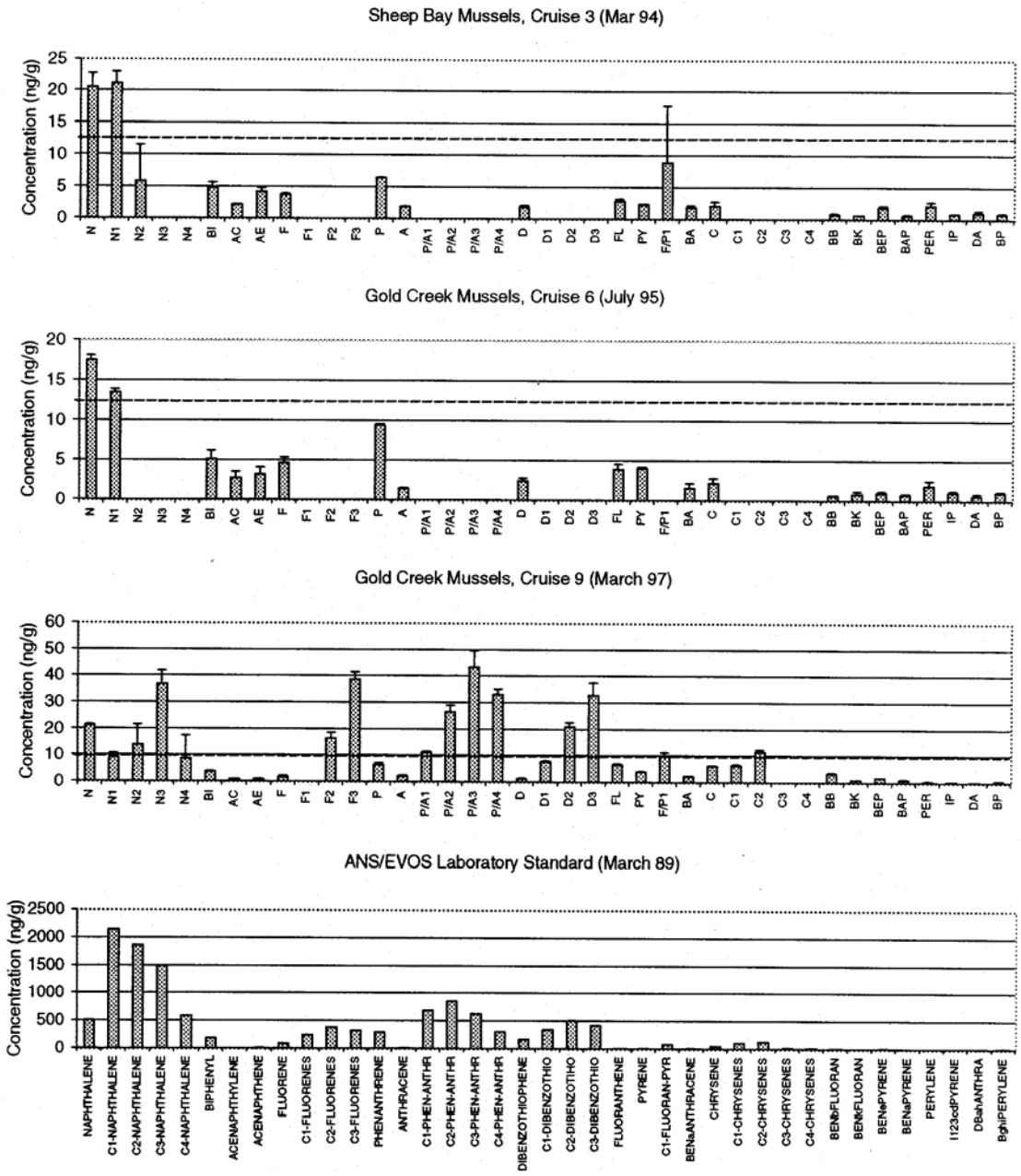


Figure 3-2. Histogram plots showing the observed distribution of 39 target PAH constituents in representative mussel tissue samples and fresh ANS/EVOS oil. Concentrations are in ng/g dry weight of tissue or oil extracted (note the different scales). The histograms for the mussel tissue samples represent the mean of three replicate analyses, and the "error bars" shown above each component represent the standard error of the arithmetic mean for each measurement. The dashed horizontal line in the plots represents the average method detection limit (MDL) for the individual PAH in a sample.

Table 3-1. Hydrocarbon Parameters Used in the LTEMP Data Analysis (Adapted from KLI, 1997).

Factor	Relevance
TPAH (mussel tissue and sediments)	Total PAH as determined by high resolution GC/MS with quantification by selected ion monitoring; defined as the sum of 2 to 5-ring polycyclic aromatic hydrocarbons: Naphthalene + fluorene + dibenzothiophene + phenanthrene + chrysene, and their alkyl homologues + other PAHs (excluding perylene); useful for determining TPAH contamination and the relative contribution of petrogenic, pyrogenic, and diagenic sources
FFPI (mussel tissue and sediments)	The Fossil Fuel Pollution Index is the ratio of fossil-derived PAHs to TPAH and is defined as follows: $FFPI = (N + F + P + D)/TPAH \times 100$ where: N (Naphthalene Series) = C_0-N + C_1-N + C_2-N + C_3-N + C_4-N F (Fluorene series) = C_0-F + C_1-F + C_2-F + C_3-F P (Phenanthrene/Anthracene series) = C_0-A +C_0-P + C_1-P + C_2-P + C_3-P + C_4-P D (Dibenzothiophene Series) = C_0-D + C_1-D + C_2-D + C_3-D FFPI is near 100 for petrogenic PAH; FFPI for pyrogenic PAH is near 0 (Boehm and Farrington, 1984)
TAHC (sediments)	Total aliphatic hydrocarbons quantifies the total n-alkanes (n-C_{10} to n-C_{34}) + pristane and phytane; represents the total resolved hydrocarbons as determined by high resolution gas chromatography with flame ionization detection (CIC/FID); includes both petrogenic and biogenic sources
UCM (sediments)	Petroleum compounds represented by the total resolved plus unresolved area minus the total area of all peaks that have been integrated; a characteristic of some fresh oils and most weathered oils
CPI (sediments)	The carbon preference index represents the relative amounts of odd and even chain alkanes within a specific boiling range and is defined as follows: $CPI = 2(C_{27} + C_{29})/(C_{26} + 2C_{28} + C_{30})$ Odd and even numbered n-alkanes are equally abundant in petroleum but have an odd numbered preference in biological material; a CPI close to 1 is an indication of petroleum and higher values indicate biogenic input (Farrington and Tripp, 1977)
CRUDE Index (sediments)	A summation of TPAH, TAHC and UCM weighted to assess the petrogenic fractions $CRUDE = (TPAH \times FFPI/100) + (TAHC/CPI^2) + UCM/1000$
MPI (mussel tissues)	The Mytilus Petrogenic index isolates the FFPI fraction of TPAH (same as first term in CRUDE) $MPI = TPAH \times FFPI/100$

In examining the oil chemistry data for this program, we initially used a fairly standard approach based on historical precedent and experience for analyzing the data. Initially,

the project chemist reviewed the common indices of total PAH (TPAH), total AHC (TAHC), UCMs, and ratio indices such as FFPI and CPI. Then, detailed studies were undertaken on the aliphatic and aromatic hydrocarbon histogram profiles generated from the triplicate analyses of sediment and tissue samples for each station and cruise. Detailed narratives from these findings are presented on a station-specific basis in Appendix I. These visual data analyses along with several new methods presented in the following sections were then integrated with the results from several other double ratio approaches to produce the final assessments on each suite of samples from each station over time.

3.1.2 The CRUDE Index

In this study, we developed the CRUDE index to encompass many of the indices presented in Table 3-1 into a single value that indicates the likely presence of crude oil. It is essentially an empirical weighted summation with emphasis towards petrogenic indicators.

$$CRUDE = (TPAH \times FFPI/100) + (TAHC / CPI^2) + UCM/1000$$

Recall that the FFPI is roughly the proportion of selected compounds in TPAH that tend to be present in petrogenic rather than pyrogenic sources, i.e., as FFPI approaches 100, it confirms a petrogenic source.

Therefore, the first term in the CRUDE index uses the FFPI/100 to isolate the petrogenic fraction of TPAHs. In a similar fashion, the second term uses the CPI to distinguish between petrogenic and biogenic sources of aliphatic hydrocarbons. A higher CPI indicates a biological source of aliphatics, so in the second term, a higher CPI lessens the contribution of TAHC to the CRUDE index value. Because TAHC is often larger than TPAH, the CPI is squared to de-emphasize this component. On rare occasions with fresh oils, the CPI may become less than 1.0 which, when squared, affects the TAHC term in an unintended manner. In these cases (as happened only twice in this study), we adjusted the CPI to 1.0. The final term of the CRUDE Index is the unresolved complex mixture (UCM), an enigmatic hump observed on chromatograms of highly weathered and polar products that cannot be adequately separated. The UCM is typically so large that it is measured in a higher magnitude of weight. To de-emphasize its mass, it is divided by 1000 (other formulations of CRUDE left it out entirely but dividing seemed to work best).

The CRUDE index appears to be useful in summarizing the five commonly used indices, and it truly aids in tracking the probable presence of petrogenic hydrocarbons. It should be noted, however, that it is a somewhat subjective *empirical* index, and that the CRUDE index values are not directly comparable to total hydrocarbon loadings, such as TPAH or TAHC. It is extremely useful in a relative sense, and we used the CRUDE index values extensively during the data analyses in this program to highlight the presence of crude oil and track changes in relative hydrocarbon concentrations over time. Although we found the CRUDE index to be a useful monitoring tool, it could not be blindly used as a

replacement for the arduous task of fully evaluating all of the available data. Therefore, as shown by the individual site-specific narratives in Appendix I, all available data were evaluated before reaching any conclusions regarding the presence or absence of ANS crude oil or EVOS residues in any sample.

3.2 Mytilus Petrogenic Index

Similar to the rationale for creating the sediment CRUDE index (see above), the *Mytilus* Petrogenic Index (MPI) was developed for tissues. It simply uses the sum of the fossil fuel pollution index (FFPI) compounds rather than TPAH to track the probable petrogenic PAH levels. In another perspective, MPI is simply the first term of the CRUDE index. This modification was required because there were no aliphatic data for the mussel tissue analyses (due to matrix interference). At sites dominated with EVOS residues, similar results were obtained with either MPI or TPAH tracking. At these sites, the MPI acted as a smoothing function by plotting just the FFPI proportion of the TPAH. However, at a site with lower FFPI proportions, the MPI tracks the petrogenic fraction better than the TPAH values, which may also include pyrogenic constituents.

3.3 MOPI

In our proposal for this program, we suggested that the Marine Oil Pollution Index (MOPI) (Payne et al, 1985) may be useful in assessing oiling loads in the RCAC samples. Unfortunately, the GERG protocols for analyzing the samples did not include determination of total resolved aliphatic hydrocarbons beyond the individual n-alkanes from n-C10 to n-C34 plus pristane and phytane. In addition, their approach does not include the discrimination of UCM fractions needed to compute MOPI. We attempted to use the index without the full data required, but abandoned the effort. Without the additional discriminating data, the partial MOPI values correlated with UCM at 0.95. As such, it provided little new useful information.

3.4 Short & Heintz Model for EVOS Weathering

Short and Heintz (1997) present an innovative model to assess PAH results for the presence of *Exxon Valdez* oil. The model is able to calculate out the confounding effects of weathering and give the statistical probability that a sample represents *Exxon Valdez* oil or natural background PAH. Unfortunately, due to nonlinear weathering behavior at the low PAH concentrations measured in this program, and issues relating to precision near and below method detection limits (MDLs), the model could not be used with most of the RCAC samples. Only 34 of the 300 sediment replicates would meet the model's criteria as stated by its developers.

Addendum – Just before final report puiblication, Jeff Short graciously ran the LTEMP data through his model after loosening the model's MDL and TPAH rejection criteria. The unqualified results are presented in Appendix VIII showing degree of weathering and degree of difference (MSE) to fitting either an EVOS or background (Constantine Harbor) signature. The results from the 34 replicates which did meet initial model criteria generally confirm our reported detections of ANS and background sources.

3.5 Short & Babcock Criteria for EVOS Versus Background Oil Signals

In their 1996 paper, Short and Babcock suggest fingerprinting criteria for screening for EVOS crude.

> "In sediment and mussel samples, the abundance of the alkyl-dibenzothiophene homologs relative to the alkyl-phenanthrenes was used to distinguish *Exxon Valdez* oil from other crude oils (or coals), and the abundance of the alkyl-chrysenes relative to the alkyl-phenanthrenes was used to distinguish crude oils from products refined from crude oils, as follows: The pattern of PAH concentrations in the samples was judged similar to that in *Exxon Valdez* oil (and is denoted as EVO-PAH) if each of three criteria was consistently met in all replicated samples, such that (1) the ratio of alkyl-dibenzothiophene homologs (summed) to alkyl-phenanthrenes (summed) exceeded 0.29, (2) the ratio of alkyl-chrysenes (summed) to alkyl-phenanthrenes (summed) exceeded 0.05, and (3) the concentration of alkyl-phenanthrenes (summed) exceeded 50 nanograms per gram (ng/g) dry weight in mussels or 20 ng/g dry weight in sediments (this latter criteria is necessary to ensure that chrysenes would be detected if present)."

As a screening tool, we built these criteria into an Excel spreadsheet and used it to define "hits" for initial evaluation of the database.

Short and Babcock (1996) also present data on the relative ratios of several other summed alkylated PAH (naphthalenes, fluorenes, phenanthrenes, dibenzothiophenes, and chrysenes) in samples that were characterized as being representative of the background "petrogenic" signal observed elsewhere throughout Prince William Sound. Based on available information for these constituents in Katalla seep oil and other sources of background hydrocarbons introduced to the east of Prince William Sound (Short and Babcock, 1996, Page et al., 1995, 1996, Bence and Burns, 1995), we used the following bracketed values for these ratios to further differentiate background signals from EVOS or ANS crude oil. For a sample to be characterized as "background", the following ranges of values had to be observed for three specific ratio-pairs in the final reduced data:

Summed alkylated naphthalenes/summed alkylated phenanthrenes = 0.65-1.0

Summed alkylated fluorenes/summed alkylated chrysenes = 1.0-1.2

Summed alkylated dibenzothiophenes/summed alkylated naphthalenes $\leq$ 0.25.

The data for each individual constituent in each sample (three replicates) at each station and for each sampling period were printed as a histogram for detailed examination by the project chemist. Mean values and standard error bars (of the arithmetic mean) for each analyte in each sample were then visually examined and manually correlated with the ratio values noted above (as a screening tool) to help differentiate ANS crude oil or EVOS residues from other sources. Detailed narratives that describe the results of these analyses are presented for each site in Appendix I.

3.6 Page et al. Calculation of Percent EVOS Oil

In a concept similar to Short and Babcock, Page et al. (1995) published a formula based on C2-phenanthrene, C2-dibenzothiophene and TPAH ratios to calculate the total amount of *Exxon Valdez* oil in a given PWS sample. This formula was also built into an Excel spreadsheet and used to evaluate the data base.

Percent EVOS = 100 x (C2Phen/0.93) x (1- (1.07 x (C2DBT/C2Phen))/0.92)/TPAH

In completing our source analyses of the data, we first used the Short and Babcock criteria to identify whether or not EVOS- or ANS-related oil was present in any given sample. Then, if the sample was classified as a "hit", the Page et al. formulation was used to calculate the relative percent contribution of EVOS or ANS oil to the overall TPAH burden identified in the sample.

In his review comments for this study, Short criticizes the foundations of Page's model and its application to the RCAC data set:

> "This model has some awkward characteristics that have not been emphasized by its developers, such as predictions of source contributions that are negative. It is also non-conservative, in that partitioned contributions may sum to more than the TPAH present. The application of this model to the subject [RCAC] data set requires a considerable extrapolation, in that the model is based on analysis of intertidal sediments that were far more contaminated than those analyzed for this program. Close inspection of the data on which the model is based reveals a quadratic dependence of the model parameters that are used to distinguish ANS from the regional PAH source (i.e. C2D and C2P), and that the parameters for these sources progressively converge at lower TPAH. The model consequently underestimates ANS contributions at lower TPAH concentrations. Also, the precision estimates for the model parameters are derived from samples that were often pseudoreplicated, and that contained very high PAH concentrations. The actual precision is much lower for PAH concentration measurements that are lower by factors of 10 to 100. The combination of systematic bias and lower analytical precision at lower TPAH concentrations accounts for most of the problems associated with this model when extrapolated below its validated range, as must be done here if it is to be applied."

We acknowledge and accept the above criticisms; however, in this study, we are technically limited in making any adjustments to the model.

Thus, we've merely made the calculations and report the results (including the negative percentages and values greater than 100 percent with a caveat to the readers that the values may be biased and imprecise.

3.7 Definition of Hot Spot or Highly Variable Sites

After the available data for all samples at all sites were evaluated, our approach was then to identify objective quantifiable criteria from the assembled data to establish a threshold that we could use to identify certain samples as being "hot or highly variable." These

objective criteria could be not be identified *a priori*, but included potential options such as target ratios of individual or summed components, intersite variability, histogram profiles meeting certain minimum parameters, absolute TPAH or TAHC concentrations, and minimum quantitative changes in CRUDE index values, etc.

3.8 Evaluation of Depth Dependence

Because GPS positioning is used for locating the sampling stations, there is an unexpected variance in sediment sampling depths. At a few sediment sampling sites, we noted that radical shifts in grain size distributions occurred between or within cruises and, in some instances, these changes in grain size correlated with the oil indices (Table 17, KLI, 1997). Because sediment oil loading is known to be affected by grain size distribution (i.e., a function of particle size, interstitial water flux, microbial activity, etc.), any shifts in bottom type introduces variability that confounds the effort to track temporal trends. A shift in grain size suggests either a major sediment transport event (e.g., big storm waves), patchy sediment distribution, or a field positioning issue. To assess the effect, we plotted the six oil index values with the grain size, TOC and TIC data along with the reported sampling depth for each replicate to see if any trends were apparent.

3.9 Data Manipulation and Statistical Analyses

3.9.1 Database Manipulation and Formatting

The RCAC oil chemistry database was received directly from KLI as files attached to e-mail. Conversion into an Rbase database was accomplished with minimal difficulty. Most effort was involved with creating logical subsets of data and reassigning field names to preferred nomenclature. Copies of the pertinent tables were exported to Microsoft Excel and e-mailed to the other principal investigators for the data manipulations and detailed evaluation. Because all transfers were handled electronically, there was no compromise to the initial data integrity.

Other data received electronically and reviewed for this study included the EVOS Trustee Council's hydrocarbon database, spreadsheets from NOAA HAZMAT studies, data from the NOAA Mussel Watch web site, and finally, paper reports from Shaw and Feder's Port Valdez studies (Feder & Shaw 1987, 1990; Shaw et al. 1985, 1986, 1993, 1994).

3.9.2 Inferential Analyses in Other Programs

Several approaches have been used for statistical analysis of oil chemistry results from Prince William Sound. Early on the scene is Shaw's extensive work around Port Valdez using multiple comparisons between samples of each individual compound to assess for yearly changes. This approach seems adequate to tracking the few compounds of interest

for regulatory purposes, but is awkward for assessing the wide concerns of the RCAC program.

The chemistry portion of the ongoing NOAA HAZMAT study (Henry, et al. 1995, 1996) does some interesting things with multivariate analyses using some of the same fingerprint signature ratios that this review uses. However, most of their work is geared towards source identification and weathering issues at highly contaminated sites rather than inferential tests and tracking trends.

The NOAA Status and Trends program (including Mussel Watch) has only two sites in Alaska. Their statistical methods are undetermined at this time, but their samples are not directly comparable to RCAC tissues because they pool the mussels into an unreplicated composite sample. They also have indicated they prefer not to make lipid adjustments to the tissue chemistry results because that correction increases the variability of the data considerably (Tom O'Connor, *pers. comm.*).

An approach common to a variety of researchers, notably KLI and Exxon, is the classic procedures of parametric inferential testing using screening tests for data normality, equivalence of variance, and discrimination of outliers followed by a suite of transformations, ANOVAs, and *a posteriori* multiple comparisons. We take a similar approach, but instead use randomization testing techniques that do not place constraints on the data distribution (discussed below).

Both KLI and Exxon also used multivariate techniques such as clustering and ordination to gain another perspective on the data. In this study, we also used clustering and nonmetric multidimensional scaling to examine trends. However, we do not believe the results justify the complexity of presenting the material and so they are not included in this report.

To our knowledge, LTEMP has been the only study in PWS to provide a power analysis of its sampling program (KLI, 1993). In that analysis, KLI found that 3 replicates would be adequate to reach a target statistical power goal of 0.80.

3.9.3 Our Approach to Inferential Analyses

For testing significance between sites and cruises, we used TPAH, TAHC and UCM values from laboratory analyses and we calculated FFPI, CPI, and CRUDE indices as the primary data of interest (Table 3-1). The CRUDE index was developed during this study (see below). Because the mussel data have not been previously examined, all available morphometric and chemical variables were used in the initial analyses.

The data were partitioned into three primary oiling-source categories and three spatial categories. Oiling source categories comprised EVOS sites, sites with potential impact from oil transport, and reference sites. Spatial categories are obviously sites inside Port Valdez, inside Prince William Sound, and the Gulf of Alaska sites.

In previous reports, KLI used a variety of parametric statistics (normality tests, outlier tests, ANOVAs, and t-tests) to produce their results. This approach, while perfectly valid, expends a lot of effort ensuring that the data fit normal distributions prior to testing for significant effects. We prefer to use randomization statistics (sometimes called exact or resampling statistics) which makes no assumption regarding the data distribution, requires no transforms, and produces an exact measure of the test's significance without resorting to tabled values. If the data being tested happened to be distributed exactly normally, then a randomization test would produce the same results as its parametric counterpart. However, when the data are not normally distributed, randomization statistics produce an exact probability estimate without the bias from approximations associated with parametric methods (particularly with small sample sizes). The original method was first suggested in 1928 by the father of modern statistics, Karl Fisher, who acknowledged the required computational impracticality prior to the advent of computers. This category of statistics includes Monte Carlo simulation, jack-knife, boot-strap, and randomization tests; we used the latter. For comparison testing, we used either two-tailed, two sample t tests or one-way ANOVAs.

One of the benefits of randomization testing is the ability to design custom analyses. For this study, a trends test was specially designed to assess whether a time series of data (all cruises at a site) showed more of a trend than you might expect from a set of randomized data. The concept involves developing a trend index value for the original time-series data set and then successively rescrambling (randomly selecting from the overall pool of values) the time series and recalculating the index, and then see if its value was higher than the original (i.e., representing a more ordered series). The scrambling and assessment is redone thousands of times (the randomization part), and where the initial index value falls in the final distribution of scrambled results defines the exact probability of the original data set being different from random. The trend indexing procedure simply encodes each sampling event into +1 when values increase from the previous sampling and –1 when they decrease. The absolute sum of the values becomes the index (Driskell et al., in prep). Predictably, a random data set will tend to have low index values (the pluses and minuses cancel out) while a more ordered series will have a higher index value (up to n–1) that peaks when the data are uniformly increasing or decreasing.

Randomization ANOVA and t tests were programmed in Quick Basic using modified code from Edgington (1987). The randomization time series test was programmed in Visual Basic for an Excel spreadsheet.

Power analyses were calculated for all sediment and tissue samplings to determine whether 3 replicates were still sufficient for attaining the target power of 0.80. Using log or arcsine transformed means, an alpha of 0.05 and assuming equal variance for a two-sample t test, the power of detecting a 25-, 50-, 75-, or 100-percent change in the mean of each oil index was calculated for every station and cruise. Power statistics were computed using Excel add-in power functions from PIFACE software.

4. Results and Discussion

4.1 Temporal and Geographical Variation in Hydrocarbon Signatures and Sources

4.1.1 Problems with Non-continuous Sampling and Site Coverage

In the 4.5 years of data collection during this program, there have been some changes in the sampling design regarding station coverage for deep and shallow sediments. However, mussel tissue collections have been essentially continuous for the March and July samplings at all stations for the entire period.

Deep sediment samples were collected at most stations through cruise 4 (July 1994) and then the sampling plan was changed to July samplings only (except in Port Valdez, which stayed on biannual sampling). At the same time, shallow sediment stations were added to most sites in the program on a biannual sampling basis, which has continued through July 1997 (the end of available data). Table 4-1 presents a summary of sediment sample coverage over the duration of the program. Unfortunately, while this schedule appears to have been a beneficial adjustment to the program scope, analyses in this report are compromised by the discontinuities.

Table 4-1 - Sediment sampling plan overview, visits by depth and cruise.

		Cruise No.									Total
		1	2	3	4	5	6	7	8	9	Site
Station	Dept	Mar-93	Jul-93	Mar-94	Jul-94	Mar-95	Jul-95	Mar-96	Jul-96	Mar-97	Visits
AIB	S	x	x	x	x		x		x		6
AMT	S	x	x	x	x	x	x	x	x	x	9
DII	M				x	x	x	x	x	x	6
	S	x	x	x	x		x		x		6
GOC	S	x	x	x	x	x	x	x	x	x	9
KNH	A		x	x	x		x		x		5
	M					x	x	x	x	x	5
	S	x									1
SHB	M				x	x	x	x	x	x	6
	S	x	x	x	x		x		x		6
SHH	M					x	x	x	x	x	5
	S		x	x	x		x		x		5
SLB	M				x	x	x	x	x	x	6
	S	x	x	x	x		x		x		6
WIB	M					x	x	x	x	x	5
	S	x	x	x	x		x		x		6
Cruise Totals		8	9	9	12	8	15	8	15	8	92

Depth Keys M = shallow (mid) depth sediments
S = deep depth sediments
A = tanker anchorage sediments

An additional complication is that there is little in common between the deep and shallow sediment stations at any one site. The energy regimes are considerably different, as reflected by differences in grain size distributions between the deep and shallow stations. The depositional regime (currents, topography, and particulate sources) affects the accumulation of organics, particularly biogenic and petrogenic hydrocarbons that tend to associate with finer grained particulates. Because the deep and shallow depositional regimes are not comparable, a continuous sediment record is not available for analysis except at Alyeska Marine Terminal (AMT) and Gold Creek (GOC). In other words, the abbreviated temporal trends can be evaluated separately for the different depths at any given station, but the sample sizes are small and results from different depths cannot generally be readily compared. As discussed in detail for each station in Appendix I, the absolute differences in TPAH and TAHC levels between the deep and shallow sediments at any given station can vary by as much as an order of magnitude, and in terms of composition, the PAH and AHC profiles are also very different. That is, the PAH pattern is not always parallel in deep and shallow sediments.

4.2 Sediment Data Evaluation

4.2.1 Overall Concentrations and General Sources

4.2.1.1 Total Polynuclear Aromatic Hydrocarbon (TPAH) and Total Aliphatic Hydrocarbon (TAHC) Burdens

TPAH and TAHC concentrations in both deep and shallow sediments at all stations are generally very low (Table 4-2). TPAH concentrations in the replicated deep sediment samples range from less than 30 ng/g dry weight in the cleaner sediments at Aialik Bay to over 1,670 ng/g dry weight in Sleepy Bay. Shallow sediments have generally even lower TPAH levels, ranging from less than 10 ng/g dry weight at Knowles Head to approximately 400 ng/g dry weight at Sleepy Bay.

TAHC concentrations in the deep sediments range from around 300 ng/g dry weight in the Knowles Head anchorage region to over 2,500 ng/g dry weight in Windy Bay. With the exception of the deep sediments at Alyeska Marine Terminal in Port Valdez, however, almost all of these aliphatic hydrocarbon loadings are derived from biogenic sources. Shallow sediments have generally lower TAHC levels, ranging from less than 50 ng/g dry weight at Knowles Head to around 1,690 ng/g dry weight at Sleepy Bay (which is in fact much higher than all the other stations).

Table 4-2. Mean Sediment Results by Station and Survey, Cruises 1-10, Mar 93-Jul 97.

StnId	CRUDE (wt sum)	TPAH (ng/g)	TAHC (ng/g)	UCM (ug/g)	FFPI (ratio)	CPI (ratio)	MOPI (sum)	TOC (%)	Sand (%)	Silt+Clay (%)	Average Depth (m)	Sampling Date
AIB-S-1	16.9	11.6	233	0.0	87.9	7.1	-1.4	0.7	22.4	77.6	-28.4	Mar-93
AIB-S-2	37.8	36.4	305	0.6	83.2	6.6	1.2	0.7	14.7	85.3	-37.8	Jul-93
AIB-S-3	29.6	29.3	245	5.1	80.2	16.3	5.1	0.4	19.0	81.1	-35.4	Mar-94
AIB-S-4	31.3	35.8	248	0.0	80.4	11.2	-0.9	0.6	24.8	75.3	-25.6	Jul-94
AIB-S-6	14.8	15.8	217	0.8	73.2	9.2	0.7	0.6	33.1	66.9	-30.6	Jul-95
AIB-S-8	37.9	25.4	203	6.3	81.3	12.2	4.1	0.6	28.6	71.4	-33.6	Jul-96
AMT-S-1	1181.3	242.6	2091	122.2	60.8	1.5	9.7	0.8	7.4	92.6	-79.2	Mar-93
AMT-S-2	1440.3	246.0	2018	120.6	56.4	1.3	9.6	0.7	5.6	94.4	-69.9	Jul-93
AMT-S-3	489.7	202.5	1473	98.8	53.9	2.3	9.4	0.6	5.7	94.3	-66.6	Mar-94
AMT-S-4	680.4	264.4	1530	93.2	57.9	1.9	9.2	0.7	4.3	95.7	-73.0	Jul-94
AMT-S-5	738.6	212.0	1390	98.7	45.7	1.6	9.5	0.6	5.2	94.9	-70.3	Mar-95
AMT-S-6	2234.3	880.2	2275	134.2	62.9	1.2	9.7	0.8	4.9	95.1	-71.2	Jul-95
AMT-S-7	365.3	201.8	1262	101.8	57.9	3.1	9.5	0.5	2.9	97.1	-70.6	Mar-96
AMT-S-8	592.9	302.5	1883	108.5	62.3	2.5	9.4	0.7	4.4	95.6	-71.3	Jul-96
AMT-S-9	762.7	417.8	2370	1.0	63.0	2.3	2.1	0.8	7.3	92.8	-66.8	Mar-97
AMT-S-10		303.2	1498	89.6								Jul-97
DII-M-4	17.0	22.0	103	3.6	54.4	7.2	5.1	0.3	94.6	5.4	-6.7	Jul-94
DII-M-5	18.2	15.3	93	3.3	63.4	4.5	5.2	0.2	98.2	1.8	-6.3	Mar-95
DII-M-6	10.8	9.6	97	0.3	41.6	3.7	-0.4	0.2	98.1	1.9	-7.1	Jul-95
DII-M-7	44.2	16.1	143	17.6	48.3	2.8	8.1	0.2	94.6	5.4	-5.8	Mar-96
DII-M-8	70.5	24.1	156	12.5	73.4	3.6	7.4	0.4	96.7	3.4	-5.3	Jul-96
DII-M-9	55.6	23.0	161	1.9	66.4	2.2	3.7	0.4	97.0	3.0	-5.9	Mar-97
DII-M-10		25.4	136	4.2								Jul-97
DII-S-1	287.6	165.6	740	0.4	72.9	2.2	0.9	0.4	70.1	29.9	-32.4	Mar-93
DII-S-2	401.6	325.1	813	15.2	65.3	2.1	6.3	0.8	61.2	38.8	-33.9	Jul-93
DII-S-3	223.4	202.3	580	21.4	67.6	3.1	7.2	1.1	64.6	35.4	-38.3	Mar-94
DII-S-4	210.6	186.3	560	10.5	73.5	3	5.9	1.1	70.7	29.2	-31.2	Jul-94
DII-S-6	231.9	150.0	442	5.0	75.7	2	4.7	1.0	69.5	30.5	-31.0	Jul-95
DII-S-8	272.2	201.7	599	93.0	72.8	4.5	9.6	1.0	71.5	28.5	-39.2	Jul-96
GOC-S-1	38.2	47.3	946	6.2	61.0	15.9	4.1	0.7	20.6	79.4	-30.6	Mar-93
GOC-S-2	29.4	37.7	567	3.7	58.5	12.1	3.7	0.6	11.5	88.5	-29.3	Jul-93
GOC-S-3	48.4	70.6	879	3.3	59.2	14.1	2.9	0.5	11.2	88.8	-30.4	Mar-94
GOC-S-4	28.1	44.4	500	2.7	55.4	18.8	3.2	0.6	24.5	75.5	-24.3	Jul-94
GOC-S-5	21.8	40.6	438	0.7	50.9	18.5	0.7	0.6	18.4	81.6	-31.6	Mar-95
GOC-S-6	36.2	52.1	597	4.2	53.2	13.1	3.8	0.7	13.6	86.4	-27.8	Jul-95
GOC-S-7	49.2	89.1	527	14.3	40.5	14.7	6.0	0.5	12.0	88.0	-37.9	Mar-96
GOC-S-8	46.8	51.1	537	13.1	61.8	39.5	5.7	0.6	25.2	74.8	-27.9	Jul-96
GOC-S-9	37.7	44.1	499	1.7	63.1	7.9	2.4	0.7	18.4	81.7	-27.6	Mar-97
GOC-S-10		55.7	618	18.3								Jul-97
KNH-A-2	172.3	125.4	510	0.6	84.8	3	1.1	0.4	65.0	35.0	-31.8	Jul-93
KNH-A-3	169.1	260.8	214	4.5	62.2	3.4	5.3	0.5	71.9	28.1	-32.1	Mar-94
KNH-A-4	71.3	69.3	174	9.3	78.5	5	6.4	0.9	61.6	38.4	-32.0	Jul-94
KNH-A-6	62.7	66.8	161	0.0	71.0	3.2	-1.1	0.3	73.5	26.6	-35.4	Jul-95
KNH-A-8	201.5	159.1	492	3.1	80.7	2.7	3.8	0.7	53.2	46.9	-32.9	Jul-96
KNH-M-5	17.3	6.9	30	12.0	81.4	17.6	8.4	0.2	97.7	2.4	-5.6	Mar-95
KNH-M-6	14.8	6.8	47	0.1	71.2	2.2	-0.9	0.2	98.6	1.4	-6.1	Jul-95
KNH-M-7	25.4	8.6	49	0.0	83.5	1.9	-1.6	0.2	94.4	5.6	-6.0	Mar-96
KNH-M-8	9.8	9.0	35	0.0	77.4	5.6	-3.0	0.2	98.3	1.7	-5.4	Jul-96
KNH-M-9	11.0	6.3	66	0.3	59.7	3.2	1.4	0.3	97.2	2.8	-8.9	Mar-97
KNH-M-10		10.6	51	11.1								Jul-97
KNH-S-1	27.9	4.4	77	0.4	77.5	1.8	0.0	0.2	94.5	5.5	-5.4	Mar-93
SHB-M-4	30.4	37.2	167	1.3	74.7	11.4	2.8	0.9	90.0	10.0	-6.2	Jul-94
SHB-M-5	45.9	46.5	159	8.3	67.7	8.7	6.0	0.8	94.4	5.7	-5.8	Mar-95
SHB-M-6	61.0	56.6	254	1.1	58.0	3.1	1.1	0.9	94.7	5.2	-6.5	Jul-95
SHB-M-7	60.5	52.9	197	6.0	75.3	4.7	5.6	0.7	93.6	6.4	-6.7	Mar-96
SHB-M-8	119.9	70.5	392	3.2	76.0	2.7	3.8	0.7	89.9	10.1	-6.3	Jul-96
SHB-M-9	248.9	119.9	685	4.2	73.3	2.4	3.9	1.5	70.4	29.5	-9.3	Mar-97
SHB-M-10		106.0	587	13.8								Jul-97

StnId	CRUDE (wt sum)	TPAH (ng/g)	TAHC (ng/g)	UCM (ug/g)	FFPI (ratio)	CPI (ratio)	MOPI (sum)	TOC (%)	Sand (%)	Silt+Clay (%)	Average Depth (m)	Sampling Date
SHB-S-1	138.5	123.2	442	2.9	76.5	3.7	3.7	1.3	80.5	19.5	-28.6	Mar-93
SHB-S-2	104.6	55.5	321	4.5	75.7	2.9	4.7	1.8	69.5	30.5	-27.4	Jul-93
SHB-S-3	78.8	96.0	212	4.8	72.2	6.8	5.1	0.9	78.8	21.1	-33.6	Mar-94
SHB-S-4	84.0	98.0	279	2.4	75.6	6	3.6	0.9	73.7	26.3	-32.4	Jul-94
SHB-S-6	177.7	143.6	599	0.6	72.9	2.9	0.6	1.1	63.2	36.8	-34.4	Jul-95
SHB-S-8	401.7	196.8	736	1.8	80.0	2.9	2.7	1.2	34.9	65.1	-35.5	Jul-96
SHH-M-5	55.8	80.6	402	2.3	59.5	7.4	3.0	0.9	84.1	15.9	-6.3	Mar-95
SHH-M-6	135.6	48.4	251	7.1	58.1	3.7	3.9	0.3	94.4	5.6	-8.1	Jul-95
SHH-M-7	63.7	61.7	274	2.3	68.2	3.8	3.2	0.9	96.3	3.8	-6.7	Mar-96
SHH-M-8	84.1	60.5	350	5.3	69.3	3.4	4.9	0.9	83.2	16.8	-7.4	Jul-96
SHH-M-9	113.3	127.8	351	1.2	66.5	3.9	2.3	1.2	83.7	16.3	-6.1	Mar-97
SHH-M-10		73.7	251	2.4								Jul-97
SHH-S-2	162.8	279.4	1037	7.5	44.7	5.7	4.6	1.0	64.3	35.7	-30.0	Jul-93
SHH-S-3	142.1	218.0	547	9.4	56.9	8.6	4.1	1.4	65.9	34.1	-30.0	Mar-94
SHH-S-4	363.4	239.0	534	2.4	50.9	5.5	3.2	1.1	68.7	31.4	-25.5	Jul-94
SHH-S-6	163.6	241.6	764	0.2	54.6	5	0.4	1.1	67.0	33.0	-28.5	Jul-95
SHH-S-8	142.9	219.8	584	0.0	58.0	6.7	-0.3	1.5	55.9	44.1	-30.5	Jul-96
SLB-M-4	204.9	252.1	384	32.8	43.9	2.4	8.6	1.0	88.4	11.7	-7.3	Jul-94
SLB-M-5	482.9	664.5	688	81.0	30.7	1.9	9.7	1.3	88.0	12.1	-8.0	Mar-95
SLB-M-6	724.4	537.1	1693	118.8	33.9	2	9.8	2.2	88.6	11.4	-9.8	Jul-95
SLB-M-7	296.5	427.3	934	121.6	35.5	7.9	10.0	1.3	84.8	15.2	-6.7	Mar-96
SLB-M-8	257.4	481.6	1020	74.2	33.1	6.8	8.9	1.5	90.8	9.2	-6.9	Jul-96
SLB-M-9	68.0	109.7	218	10.6	43.1	5.4	6.6	0.8	95.3	4.7	-8.3	Mar-97
SLB-M-10		41.2	130	8.3								Jul-97
SLB-S-1	162.0	305.5	395	12.9	39.9	3.8	6.5	0.8	89.8	10.2	-34.3	Mar-93
SLB-S-2	143.7	197.8	262	7.5	52.2	2.8	6.0	0.6	87.7	12.3	-34.6	Jul-93
SLB-S-3	142.2	229.7	399	21.5	45.4	4.5	7.5	1.3	88.2	11.9	-31.4	Mar-94
SLB-S-4	993.3	1672.0	417	12.3	44.2	1.5	6.6	0.9	91.7	8.3	-31.6	Jul-94
SLB-S-6	156.7	256.0	280	5.6	43.4	2.5	5.4	1.0	87.2	12.8	-33.9	Jul-95
SLB-S-8	237.2	403.4	488	16.1	48.6	4.9	6.8	1.1	87.5	12.5	-31.2	Jul-96
WIB-M-5	11.2	12.8	143	0.3	78.5	12.4	-0.6	0.5	98.2	1.9	-5.5	Mar-95
WIB-M-6	7.8	7.5	232	0.0	66.0	8.2	-1.2	0.6	81.7	18.2	-5.9	Jul-95
WIB-M-7	8.6	6.5	99	0.0	81.9	6.2	-1.4	0.2	97.7	2.3	-4.6	Mar-96
WIB-M-8	5.8	8.2	155	0.1	63.0	17.3	-1.1	0.7	97.5	2.5	-5.4	Jul-96
WIB-M-9	34.0	38.2	160	0.5	66.5	8.5	1.3	0.6	94.5	5.5	-5.6	Mar-97
WIB-M-10		45.6	125	1.4								Jul-97
WIB-S-1	130.3	113.8	2126	33.4	67.5	10.5	6.4	2.4	45.3	54.7	-25.9	Mar-93
WIB-S-2	143.7	196.0	2251	11.1	61.6	12.4	4.6	3.2	36.4	63.6	-25.5	Jul-93
WIB-S-3	91.6	141.5	2180	5.2	57.7	21.2	3.3	3.0	35.0	65.0	-26.5	Mar-94
WIB-S-4	116.4	136.8	1757	7.4	73.4	14.5	3.8	2.3	44.5	55.6	-25.8	Jul-94
WIB-S-6	138.3	207.5	2528	1.2	62.7	12.4	1.8	2.7	51.1	48.9	-25.3	Jul-95
WIB-S-8	123.7	154.2	2141	5.4	71.6	16.6	3.2	2.9	31.3	68.7	-25.6	Jul-96

AIB	Aialik Bay
AMT	Alyeska Marine Terminal
DII	Disk Island
GOC	Gold Creek
KNH	Knowles Head
SHB	Sheep Bay
SHH	Shuyak Island
SLB	Sleepy Bay
WIB	Windy Bay

-M-	Mid depth sediments
-S-	Deep sediments
CRUDE	Crude oil index
TPAH	Total polycyclic aromatic hydrocarbons
TAHC	Total aliphatic hydrocarbons
UCM	Unresolved complex mixture
FFPI	Fossil fuel pollution index
CPI	Carbon preference index
MOPI	Marine oil pollution index
TOC	Total organic carbon

Caution: Cruise 10 values are early release data that have not yet been validated by KLI.

adapted from KLI 95-96 Annual Monitoring Rpt

4.2.1.2 Histogram Pattern Source Identification

With these low background levels, it is relatively easy to document any absolute increases in TPAH or TAHC concentrations from recent oil transportation activities. Perhaps as important, however, the hydrocarbon fingerprints in these sediments will also allow differentiation of new sources identified in future samplings. The data presented previously in Figures 3-1 and 3-2 illustrate this point and are discussed briefly below. More detailed discussions of the PAH and AHC histogram patterns, source identification, and analyses of changes and overall trends are presented for each station in Appendix I.

The PAH patterns for the deep sediments examined in this program reflect several "background" sources, including: the Katalla oil seeps, coal particles potentially derived from the rivers east of Prince William Sound, and oil-transportation activities associated with the Alyeska Marine Terminal in Port Valdez (Short and Babcock, 1995; Page et al., 1995, 1996, Bence and Burns, 1995). At this time, there is debate in the scientific literature as to whether the natural "background" hydrocarbons are actually derived from oil seeps and/or coal particles from outside of the Sound. However, for the purposes of this program, it is sufficient that these fingerprints can be distinguished from the patterns generated from Alaskan North Slope crude oil introduced from present-day activities or weathered EVOS residues.

The histogram plots for the sediment samples presented in Figure 3-1 illustrate several interesting features. The individual PAH concentrations for the Disk Island deep sediments are very low (generally 3-12 ng/g dry weight), and the precision as reflected by the standard error bars is very good. Several of the components are below the average individual component method detection limit (MDL) of 0.7 ng/g dry weight, but again, the precision is very tight. This pattern is classified as being from such sources as coal or seep oil from outside the Sound (Short and Babcock, 1996). The individual PAH concentrations for the Disk Island shallow sediments are approximately an order of magnitude lower than the deep sediments, and most of them are, in fact, below the average individual MDL. Nevertheless, the precision among the triplicate measurements is very good, and it is possible to identify the same general background pattern as noted for the deeper sediments. The Disk Island intertidal sediments collected during cruise 6 (July 1995) show the classic weathered PAH pattern associated with EVOS residues. The naphthalene and fluorene components have a water-washed appearance with C4-naphthalene and C3-fluorene predominating over the other alkylated homologues in each series. The phenanthrene/anthracene and dibenzothiophene patterns show significantly less weathering, and from the C2-phenanthrene/C2-dibenzothiophene ratio of 1.09 (see Section 4.2.5 below), it appears that essentially 100% of these PAHs are derived from very well weathered *Exxon Valdez* oil residues (Page et al., 1995, 1996). This is further confirmed by the alkylated chrysene series, which would be absent if a weathered diesel or other refined product was the source. In comparison, the fresh ANS/EVOS oil standard shows the alkylated naphthalenes as the most prominent constituents, with lower relative concentrations of alkylated fluorenes, phenanthrenes/anthracenes, dibenzothiophenes, and chrysenes. As noted above, with evaporation/ dissolution weathering, most of these lower-molecular-weight components are removed, leaving only

the alkylated phenanthrenes/anthracenes, dibenzothiophenes, and chrysenes at very characteristic ratios in the remaining oil residues.

The histogram plots obtained on the representative mussel tissue extracts shown in Figure 3-2 illustrates that minor concentrations of ANS oil can be readily detected. In both the Sheep Bay cruise 3 (March 94) and the Gold Creek cruise 6 (July 96) samples, most of the individual analytes are below the average individual component MDL of 12 ng/g dry weight. Nevertheless, the precision obtained on the triplicate samples is very good, and the same general pattern is obtained in both samples. In fact, this pattern is characteristic of most of the mussel samples obtained from cleaner areas throughout the study area where few, if any alkylated PAHs derived from the more common oil sources are observed. Many of the constituents (naphthalene, fluorene, phenanthrene, anthracene, fluoranthene, pyrene, and the benzo(b)fluoranthene through benzo(g,h,i)perylene suite) have been identified by NOAA (1997) as combustion-derived PAH by-products from burning oil. At the same time, Bence and Burns (1995) describe a similar profile as "procedural artifacts." In actual fact, many of these components in this exact pattern do routinely show up in procedural and field blanks analyzed during this program. They are, however, generally present at even lower concentrations.

The point to be emphasized here is that the levels shown in the top two histograms in Figure 3-2 are extremely low, and that when a pulse of oil is released, it is easily detected, as shown by the histogram obtained from the Gold Creek mussels collected during cruise 9 (March 97). In this instance the characteristic pattern of relatively fresh ANS oil can be observed and potentially traced back to the Alyeska Ballast Water Treatment Plant spill that occurred in January 1997. The oil appears to be relatively fresh, based on the relative abundance of the naphthalene components compared to the dibenzothiophenes and phenanthrenes/anthracenes, and from the C2-phenanthrene to C2-dibenzothiophene ratio, the relative contribution of PAHs from ANS crude oil to the overall PAH burden may be as high as 50%. The histogram plot also shows the presence of alkylated chrysenes, which further allows differentiation of ANS oil from a diesel source. As an additional qualifier, it should be noted that although these latter components are lower than the average reported individual PAH MDL for that cruise, the precision of their measurement is very good, they are not observed in the blanks, and the relative abundance of the C0<C1<C2 homologues suggests that they are real, and not an artifact of the measurement process.

4.2.2 Use of the CRUDE Index for Analysis of Geographical and Temporal Trends

To aid in analyzing all of the available data from this program, the CRUDE index value characterization was developed and used to evaluate spatial and temporal trends. The CRUDE index approach combines into a single value many of the numerous individual factors and characteristic ratios that had been used by chemists and environmental scientists for data analysis in the past. With this single-value approach, it is possible to plot the CRUDE index value for each station and depth over time. This facilitates the

identification of trends within a station and also allows differences to be noted among stations.

Figure 4-1 presents the CRUDE index values obtained from the sediment samples collected in the control sites at Aialik Bay, Gold Creek, and Sheep Bay; the EVOS-impacted sites at Disk Island, Shuyak Harbor, Sleepy Bay, and Windy Bay; and sites associated with tanker activities at Alyeska Marine Terminal and Knowles Head anchorage. Standard error bars reflecting the variance (of the arithmetic mean) associated with each triplicate measurement are also printed on top of each station presented in the figure. This allows an easy evaluation of apparent trends over time or among stations, with the variance associated with each measurement easily factored into the visual analysis. As noted above, sediment samples were collected at deep and shallow (mid-depth) stations. Therefore, in the figure, station identifications are denoted as DII-M-2 or DII-S-3, etc. DII-M-2 stands for DIsk Island, Mid-depth sediment, cruise 2; and DII-S-3 represents DIsk Island, deep Sediment, cruise 3, etc.

To facilitate more detailed comparisons with greater resolution on an expanded Crude Index scale, the CRUDE index plots for the deep and mid-depth stations are presented again separately in Figures 4-2 and 4-3, respectively.

As shown by the data in all three figures, relatively flat and extremely low-level CRUDE index values are obtained for the deep sediments at Aialik Bay and Gold Creek (control stations); the mid-depth sediments within Windy Bay and Disk Island (EVOS-impacted stations); and finally in the mid-depth sediments at Knowles Head (tanker-route area). At these stations, there was very little change observed in the absolute hydrocarbon concentrations, and little apparent change was noted in the patterns associated with the histogram plots generated for each station over time (see Appendix I). Likewise, these stations exhibited little or no evidence of EVOS or Alaskan North Slope crude-derived oil, and only extremely low-level background hydrocarbons from the petrogenic or coal sources outside of Prince William Sound were noted.

More significant trends in sediment hydrocarbon burdens were suggested over time by the increases in the CRUDE Index values at Sheep Bay (mid-depth and deep), at Disk Island (deep), and at Sleepy Bay (mid-depth and deep). Likewise, very high variability, and much higher absolute concentrations of petroleum-derived hydrocarbons were noted in the deep sediments at the Alyeska Marine Terminal.

Additional discussions of individual stations, and detailed correlations of histogram profiles for aliphatic and aromatic hydrocarbons with CRUDE index values, TPAH, TAHC, sample depth, grain size, etc., are presented in Appendix I.

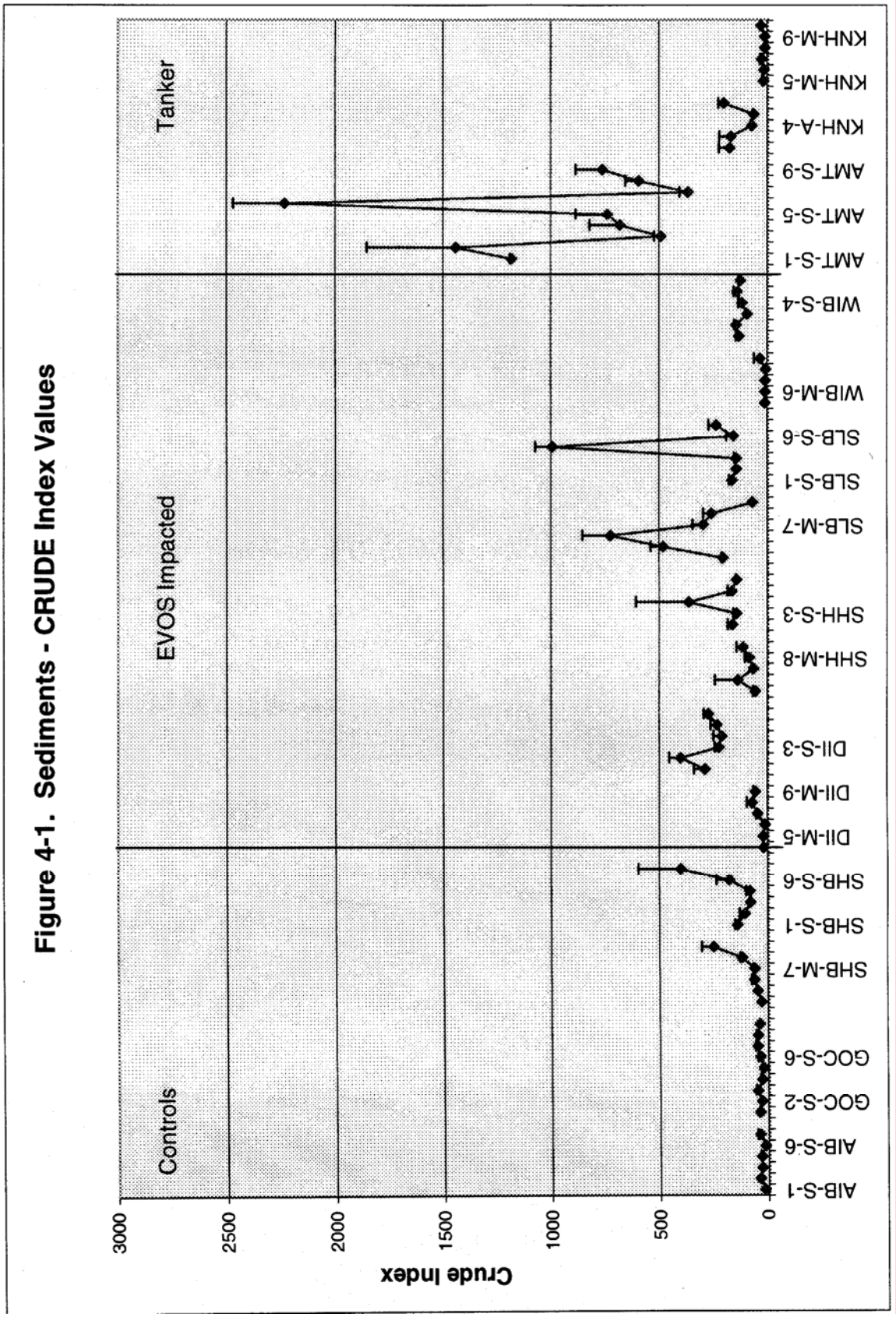


Figure 4-1. Sediments - CRUDE Index Values
Crude Index
3000
2500
2000
1500
1000
500
0
Controls
EVOS Impacted
Tanker
AIB-S-1
AIB-S-6
GOC-S-2
GOC-S-6
SHB-M-7
SHB-S-1
SHB-S-6
DII-M-5
DII-M-9
DII-S-3
SHH-M-8
SHH-S-3
SLB-M-7
SLB-S-1
SLB-S-6
WIB-M-6
WIB-S-4
AMT-S-1
AMT-S-5
AMT-S-9
KNH-A-4
KNH-M-5
KNH-M-9

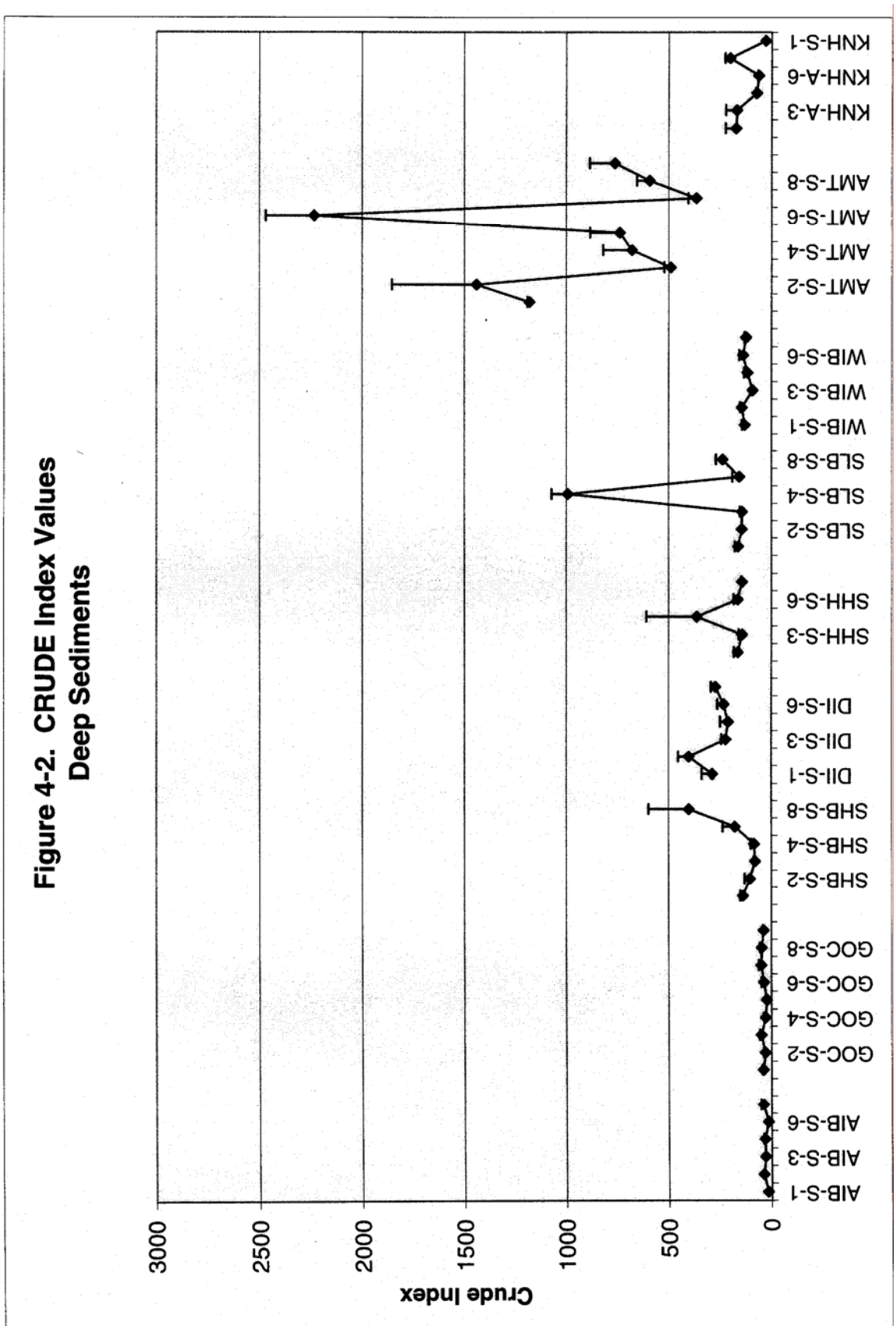


Figure 4-2. CRUDE Index Values
Deep Sediments
Crude Index
3000
2500
2000
1500
1000
500
0
AIB-S-1
AIB-S-3
AIB-S-6
GOC-S-2
GOC-S-4
GOC-S-6
GOC-S-8
SHB-S-2
SHB-S-4
SHB-S-8
DII-S-1
DII-S-3
DII-S-6
SHH-S-3
SHH-S-6
SLB-S-2
SLB-S-4
SLB-S-8
WIB-S-1
WIB-S-3
WIB-S-6
AMT-S-2
AMT-S-4
AMT-S-6
AMT-S-8
KNH-A-3
KNH-A-6
KNH-S-1

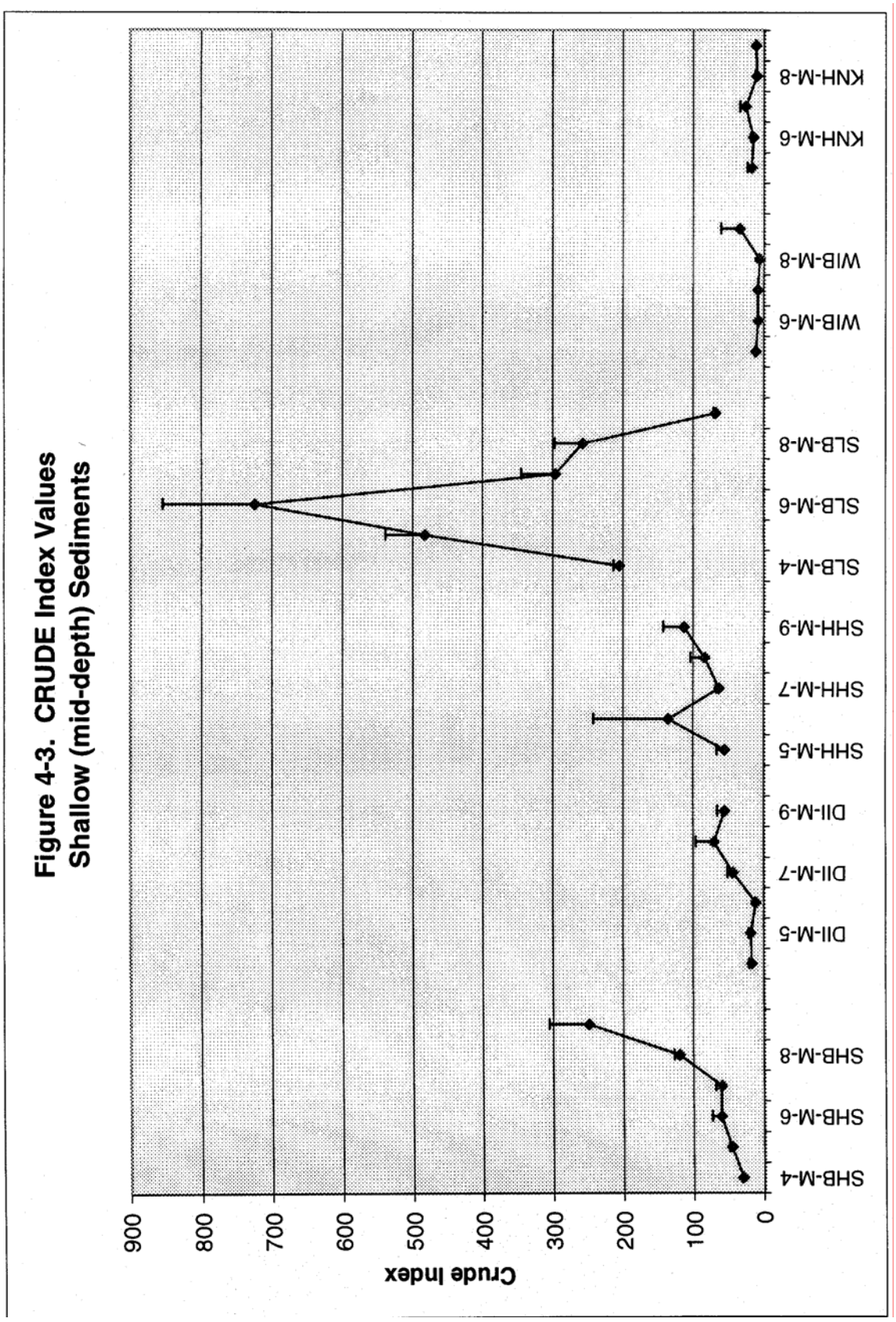


Figure 4-3. CRUDE Index Values
Shallow (mid-depth) Sediments
Crude Index
0
100
200
300
400
500
600
700
800
900
SHB-M-4
SHB-M-6
SHB-M-8
DII-M-5
DII-M-7
DII-M-9
SHH-M-5
SHH-M-7
SHH-M-9
SLB-M-4
SLB-M-6
SLB-M-8
WIB-M-6
WIB-M-8
KNH-M-6
KNH-M-8

4.2.4 Short & Babcock Criteria for EVOS Versus Background Oil Signals

As described in Section 3.5, Short and Babcock (1996) have published data that characterized the relative abundance of different alkylated PAH in background samples from Constantine Harbor sediments. They believed that the PAHs in those samples were probably derived primarily from geologic sources external to Prince William Sound and concluded that the relatively low proportions of dibenzothiophene homologues clearly indicated a source other than EVOS oil. Using their data, we defined acceptance ranges for specific ratios of summed pairs of alkylated naphthalenes, phenanthrenes, fluorenes, chrysenes, and dibenzothiophene homologues that we felt were representative of natural background signals from sources outside of Prince William Sound.

As an example of this approach, Table 4-3 presents the cruise 9 data generated for these summed ratio analyses. Appendix II presents the summed ratio analyses data for both deep and shallow sediments for all cruises.

Although not absolute, the criteria for accepting a sample as being representative of the "background" were generally as follows:

Σ Naphthalene/Σ Phenanthrene = 0.65 – 1
Σ Fluorene/Σ Chrysene = 1.0 – 1.2
Σ Dibenzothiophene/Σ Naphthalene $\leq$ 0.25

When using this approach, the histogram profiles were always examined in detail first to see if: 1) the source pattern was obvious, 2) variability among the replicates was excessive, or 3) a significant number of analytes were present at levels below the average method detection limit. Then, when the ratio data in the table were consulted, it was still often necessary to allow some leeway before categorically stating that a sample was characterized as background. For example, in some instances, not all three of the variables fell into the expected ranges, and "professional judgment" had to be exercised. Nevertheless, the quantitative data were useful as an additional screening tool when trying to visually differentiate between what otherwise appeared to be very similar profiles. The results of these analyses are discussed on a site-specific basis in Appendix I.

When attempting to positively identify a sample as being derived from ANS oil or EVOS residues, another set of ratio variables (also developed by Short and Babcock, 1996) was used. In this case, the PAH pattern was classified as a "hit" for ANS or EVOS oil if three criteria were met:

Σ Dibenzothiophenes/Σ Phenanthrenes $\geq$ 0.29
Σ Chrysenes/Σ Phenanthrenes $\geq$ 0.05
And Σ Chrysenes $\geq$ 10 ng/g dry weight

Table 4-3. Results of screening sediment PAH ratios as indicators of ANS or background hydrocarbon source in last cruise (March 97). Other cruises are presented in Appendix II.

Calculations based on Short and Babcock, 1996 and Page, et al, 1995.

			Katalla Ratios			EVOS Ratios		EVOS Eval	ANS Portion	
StnID	Rep	TPAH	N/PA	F/CHR	DBT/N	DBT/PA	Chr/PA	Hits	%ANS	Site Avg
AIB-S-8	1	28.9	1.24	3.61	0.11	0.00	0.00	na	9.3	na
AIB-S-8	2	33.4	1.38	4.69	0.07	0.00	0.00	na	11.0	
AIB-S-8	3	13.8	0.54	2.50	0.30	0.00	0.00	na	-16.5	
AMT-S-9	1	416.6	0.45	0.91	0.02	0.99	0.88	hit	55.7	54
AMT-S-9	2	449	0.39	0.61	0.02	0.86	1.05	hit	53.7	
AMT-S-9	3	387.9	0.63	0.46	0.02	0.74	0.94	hit	51.4	
DII-M-9	1	18.6	1.06	0.94	0.24	0.00	0.00	na	16.7	na
DII-M-9	2	24.6	0.91	0.67	0.21	0.00	0.00	na	15.9	
DII-M-9	3	25.7	0.96	1.58	0.16	0.00	0.00	na	8.6	
DII-S-8	1	150.1	0.95	1.20	0.02	0.00	0.00	na	7.0	na
DII-S-8	2	190.2	0.99	0.99	0.02	0.16	0.38	na	7.5	
DII-S-8	3	264.7	0.86	0.74	0.01	0.14	0.50	na	1.4	
GOC-S-9	1	53.6	0.98	1.36	0.08	0.00	0.00	na	18.6	na
GOC-S-9	2	38.6	1.19	1.51	0.09	0.00	0.00	na	10.9	
GOC-S-9	3	40.1	0.98	0.98	0.11	0.00	0.00	na	19.2	
KNH-A-8	1	148.8	1.16	1.88	0.02	0.00	0.00	na	1.9	na
KNH-A-8	2	145.4	1.17	2.04	0.02	0.00	0.00	na	3.2	
KNH-A-8	3	183	0.76	1.59	0.02	0.00	0.00	na	1.7	
KNH-M-9	1	5	2.29	2.00	0.56	0.00	0.00	na	no P2	na
KNH-M-9	2	9	5.22	6.00	0.19	0.00	0.00	na	no P2	
KNH-M-9	3	4.8	2.33	2.50	0.64	0.00	0.00	na	no P2	
SHB-M-8	1	56.4	0.87	2.85	0.07	0.00	0.00	na	6.4	na
SHB-M-8	2	62.9	0.78	3.11	0.06	0.00	0.00	na	2.7	
SHB-M-8	3	92.1	0.99	3.43	0.04	0.00	0.00	na	15.2	
SHB-M-9	1	142.5	0.96	1.06	0.02	0.00	0.00	na	1.4	na
SHB-M-9	2	90.4	1.35	1.79	0.03	0.00	0.00	na	-3.7	
SHB-M-9	3	126.8	0.53	0.50	0.04	0.00	0.00	na	-2.5	
SHB-S-8	1	163.9	1.15	2.17	0.02	0.00	0.00	na	0.2	na
SHB-S-8	2	211.2	1.12	2.13	0.01	0.00	0.00	na	1.1	
SHB-S-8	3	215.2	1.06	1.89	0.01	0.00	0.00	na	1.0	
SHH-M-9	1	31.5	0.54	0.95	0.15	0.00	0.00	na	-28.4	na
SHH-M-9	2	125.6	1.00	1.33	0.03	0.00	0.00	na	4.1	
SHH-M-9	3	226.3	1.08	0.83	0.01	0.09	0.43	na	-2.3	
SHH-S-8	1	229.8	1.03	0.82	0.02	0.10	0.41	na	-0.5	na
SHH-S-8	2	225.7	0.86	0.78	0.02	0.11	0.39	na	-0.6	
SHH-S-8	3	204	1.01	0.87	0.02	0.10	0.46	na	0.2	
SLB-M-9	1	109.6	0.45	0.43	0.10	0.21	1.19	na	5.4	11
SLB-M-9	2	126.1	0.73	0.55	0.07	0.38	1.23	hit	11.3	
SLB-M-9	3	93.5	0.59	0.71	0.08	0.02	1.14	na	-7.5	
SLB-S-8	1	481.4	0.25	0.47	0.03	0.12	0.38	na	2.5	na
SLB-S-8	2	438.9	0.28	0.45	0.02	0.13	0.46	na	2.6	
SLB-S-8	3	289.9	0.41	0.47	0.03	0.15	0.53	na	4.5	
WIB-M-9	1	100.6	2.41	3.75	0.02	0.00	0.00	na	-7.5	na
WIB-M-9	2	4.9	2.14	1.50	0.60	0.00	0.00	na	no P2	
WIB-M-9	3	9	2.85	1.33	0.24	0.00	0.00	na	no P2	
WIB-S-8	1	155.7	1.08	1.16	0.02	0.00	0.00	na	4.9	na
WIB-S-8	2	135.8	1.14	1.13	0.02	0.00	0.00	na	5.7	
WIB-S-8	3	171.1	1.26	0.91	0.02	0.00	0.00	na	49.4	

Note cautions regarding negative percent ANS calculations cited in Sections 3.5 and 3.6.

When utilizing this approach, it was necessary to examine each histogram profile to ensure that the sum of a certain group did not come from a single compound. For example, alkylated chrysenes (above a summed concentration of 10 ng/g dry weight) were a necessary component for a sample to be classified as a "hit" for ANS oil or EVOS related residues. If only chrysene (but not its alkylated homologues) was detected, then the summed variable may meet the criteria, but the sample was rejected as a hit, because the parent chrysene alone (without its alkylated homologues) would most likely be derived from combustion or other sources.

Thus, with these ratio-specific criteria, the data for each sample were examined to first see if they met the criteria to be classified as a "hit" for EVOS-derived oil, and then to see if they could be classified as "background from outside Prince William Sound." Discussions of these evaluations along with narrative description of trend analyses for each site are presented in Appendix I.

4.2.5 Page et al. Calculation of Percent EVOS Oil

If a sediment or tissue sample was classified as being a "hit" for EVOS or ANS crude oil, then the C2-dibenzothiophene/C2-phenanthrene quantification formulation of Page et al. (1995) was applied, as described in Section 3.6. This allowed us to calculate the relative percentage from ANS- or EVOS-derived PAH to the total PAH burden measured in that sample. The results of these calculations s for the Cruise 9 sediments are also presented in Table 4-3. The remaining data for all deep and shallow stations at all nine sites are presented in Appendix II. Out of over 300 sediment samples analyzed over the 4 1/2 years of the program, only 34 were identified as "hits" with the chrysene criteria set at 10 ng/g dry weight. When the chrysene criteria was lowered to 1 ng/g dry weight, the number of possible hits increased to 63. At this lower threshold level, however, the probability for incurring false positives increased dramatically, as more of the data from which these ratios were derived approached the method detection limits. For the analysis of the data discussed in this report, the more conservative chrysene cut-off criterion of 10 ng/g dry weight was used. Site specific discussions of the implications of these analyses are presented in Appendix I and are summarized in Section 5.

4.2.6 Hot Spots or Variable Sites

In order for a sample to be classified as a hot spot (area of high petroleum hydrocarbon concentration) or highly variable, several criteria had to be met.

1) The CRUDE Index value had to be a factor of 3-4 times greater than the baseline average for areas where a smooth flat signal was consistently obtained followed by a sudden rise (such as in Sheep Bay sediments during cruises 8 and 9).

2) The CRUDE Index values had to exhibit significant intercruise variability that exceeded the within-cruise variability (such as in the Alyeska Marine Terminal sediments).

3) The CRUDE Index values had to "spike" or show a consistent temporal trend suggesting an out-of-line event (such as in Shuyak Harbor shallow sediments during cruise 6 or deep sediments during cruise 4; and in Sleepy Bay shallow sediments for cruises 5, 6, 7, and 8, and deep sediments in cruise 4).

4) The ratio approach derived from Short and Babcock (see Sections 4.2.4) had to designate the sample as a "hit" with a chrysene cut-off criteria of 10 ng/g dry weight.

If all of the above absolute or objective criteria for classifying a sample as a hit were satisfied, then subjective conformation was obtained by examination of the histogram plots for the aromatic and aliphatic fractions from that station. In order for the histogram data to confirm a hit, the following additional qualitative criteria had to be satisfied:

1) The histogram profile had to visually confirm the ratio calculations by demonstrating the presence of multiple alkylated homologues for target groups.

2) Reasonable within-sample variability had to be observed by standard error bars that were less that 20 percent of the peak size for the majority of the components detected (i.e., the overall average profile or indices weren't being driven by a single outlier).

3) Multiple components were measured above the method detection limit.

4) Minimum contribution to the CRUDE Index value from exceptionally high biogenic aliphatic hydrocarbons with anomalous CPI values.

With this approach, several of the numerically generated hot spots or hits were eliminated to arrive at the final summary data on hot spots shown in Table 4-4. If all of the above criteria were satisfied, then the relative percent contribution from ANS or EVOS-derived PAHs to the total PAH burden in the sample was calculated by the C2-dibenzothiophene/ C2-phenanthrene approach of Page et al. (1995). The range of ANS or EVOS contributions to the overall TPAH burden (as a percent of TPAH) is also presented in Table 4-4 for all sites where positive hits for ANS or EVOS oil were confirmed.

Additional discussions of these results are presented in Section 5, and site-specific details are presented in Appendix I.

4.3 Mussel Data Evaluation

Program averages for PAH hydrocarbon burdens in the tissues ranged from 136 nanograms per gram (ng/g) dry weight at Shuyak Harbor to 512 ng/g dry weight at Alyeska Marine Terminal (Tables 4-5 and 4-6). Compared to analyses of mussels from the vicinity of active ports and at oiled sites following the *Exxon Valdez* oil spill, these levels are generally quite low (Houghton et al. 1992; Roberts et al. 1996).

Table 4-4. Summary of Sediment Hits and Percent ANS / EVOS Crude Contributions for Highly Variable and High Oil Concentration Samples -- LTEMP Cruises 1-9, March 93 - March 97.

Site	Depth	Temporal Changes in CRUDE Index	Overall Range of Average CRUDE Index Values		TPAH Concentration (ng/g DW)		TAHC Concentration (ng/g DW)		Cruise No. for Hot Spots or High Variability in CRUDE Index	Cruise No. for Positive Hits on ANS or EVOS oil[1]	Range of ANS or EVOS Contribution to TPAH Burden[2] (% of TPAH)		Influence of Depth Within Sampling Regime[3]	Crude Index Value from Last Cruise[4]
			Min	Max	Mean	C.V.(%)	Mean	C.V.(%)						
AIB	Deep	No	15	38	26	36	235	17		None	n.a.		CRUDE	38
AMT	Deep	Yes	365	2,230	330	66	1810	22	All (1-9)	All (1-9)	25	53.6		762
DII	Deep	No	210	402	200	33	615	21						272
	Shallow	No	11	71	18	31	125	25		None			CRUDE	56
	Intertidal				105,000	6	6360	110	6,8	6,8	92.5	100		
GOC	Deep	No	22	49	53	31	610	29		4	94.5			38
KNH	Anch.	No	63	201	136	59	310	57		None				201
	Shallow	No	10	25	8	16	45	31		None			TAHC	11
SHB	Deep	Yes	79	402	119	41	432	47	8,9	None			TPAH	402
	Shallow	Yes	30	249	64	46	309	66	8,9	None			CRUDE	249
SHH	Deep	No	142	363	240	10	693	31		3	98.7 (n=1)			143
	Shallow	Possibly	56	135	76	41	326	19	8,9	None				113
SLB	Deep	Yes	142	993	510	112	373	23	4	None				237
	Shallow	Yes	68	724	412	49	823	64	5,6,7	4,6,7	11.3	100	CRUDE	67
WIB	Deep	No	92	144	158	23	2160	11		3,4,6,8	20.5	49.4		123
	Shallow	Possibly	6	34	15	91	158	30	9	None			TAHC	34

Notes:
1) Short and Babcock (1996) multiple ratio approach (minimum total chrysene criteria set at 10 ng/g dry weight).
2) Page et al. (1995, 1996) C2-Phenanthrene/C2-Dibenzothiophene ratio approach.
3) Entry denotes possible minor confounding influence of depth on noted parameter within the respective shallow or deep sediment regime.
4) Value for last sample collected (Cruise 8 for deep sediments, and 9 for shallow sediments).

Table 4-5. Summary of Mytilus Tissue Hits and Percent ANS / EVOS Crude Contributions for Highly Variable and High Oil Concentration Samples -- LTEMP Cruises 1-9, March 93 to March 97

Site	Temporal Changes in CRUDE Index	Overall Range of Average CRUDE Index Values		TPAH Concentration (ng/g DW)		Cruise No. for Hot Spots or High Variability in CRUDE Index	Cruise No. for Positive Hits on ANS or EVOS oil[1]	Range of ANS or EVOS Contribution to TPAH Burden[2] (% of TPAH)		Crude Index Value from Last Cruise[3]
		Min	Max	Mean	C.V.(%)					
AIB	Possibly	35	231	127	61	3,8,9	None			231
AMT	Yes	56	1,500	512	89	3,4,5,9	1,3,4,5,7,8,	49	100	496
DII	Yes	55	742	250	92	3,4,5,9	3,4,5,8,9	30	82	290
GOC	Yes	49	530	386	70	1,3,4,5,9	1,3,5,9	49	71	331
SHB	Yes	48	318	178	72	2,4,8,9	None			318
SHH	Possibly	34	289	136	89	8,9				268
SLB	Yes	52	1,690	487	137	3	1,3,4,5,9	23	54	245
WIB	Yes	33	463	144	110	9	9	20		462

Notes:
1) Short and Babcock (1996) multiple ratio approach (minimum total chrysene criteria set at 10 ng/g dry weight).
2) Page et al. (1995, 1996) C2-Phenanthrene/C2-Dibenzothiophene ratio approach.
3) Value for last sample collected (Cruise 9 for Mytilus).

Table 4-6. Average Concentration of TPAH in Mussel Tissues in Samples from Specified Sites and Sampling Periods.

Average TPAH in Mussel Tissue (ng/g dry weight)

Station/Survey Period	Mar-93	Jul-93	Mar-94	Jul-94	Mar-95	Jul-95	Mar-96	Jul-96	Mar-97	Overall Average
Reference Sites										
Aialik Bay	71	105	194	126	56	55	92	151	292	127
Gold Creek	618	127	549	779	645	77	151	133	391	385
Sheep Bay	44	293	97	204	66	78	111	321	391	178
Reference Average	244	175	280	369	255	70	118	202	358	230
Std. Dev.	324	103	238	356	337	13	30	104	57	206
Coeff. of Variation (%)	133	59	85	97	132	19	26	51	16	89
EVOS Sites										
Disk Island	107	92	290	813	249	113	117	120	350	250
Shuyak Harbor	--	58	83	68	59	56	100	341	319	135
Sleepy Bay	358	92	2209	386	623	162	130	125	299	487
Windy Bay	65	84	126	86	62	53	112	149	559	144
EVOS Average	211	78	806	180	248	90	114	205	392	256
Std. Dev.	159	16	1025	348	265	52	12	106	120	384
Coeff. of Variation (%)	75	21	127	194	107	58	11	52	31	150
Tanker Sites										
Alyeska Marine Terminal	325	248	797	1581	517	87	242	229	582	512
Knowles Head	72	106	411	376	138	101	145	365	473	243
Tanker Average	199	177	604	978	327	94	193	297	527	378
Std. Dev.	126	71	193	603	190	7	48	68	55	359
Coeff. of Variation (%)	64	40	32	62	58	7	25	23	10	95
Overall Average	208	134	529	491	268	87	133	215	406	274
Std. Dev.	207	80	673	495	255	35	45	101	109	151

4.3.1 *Mytilus* Petrogenic Index

Figure 4-4 presents the data generated for the *Mytilus* Petrogenic Index, which is derived from the sums of individual aromatic compounds that are characteristic of petrogenic rather than pyrogenic sources (see Section 3.1.2). As with the CRUDE Index plot presented for the sediment analyses, standard error bars reflecting the variance associated with each triplicate measurement are also printed on top of each sample. This allows an easy evaluation of apparent trends or changes over time or differences among stations, with the overall variance associated with each measurement easily factored into the visual analysis.

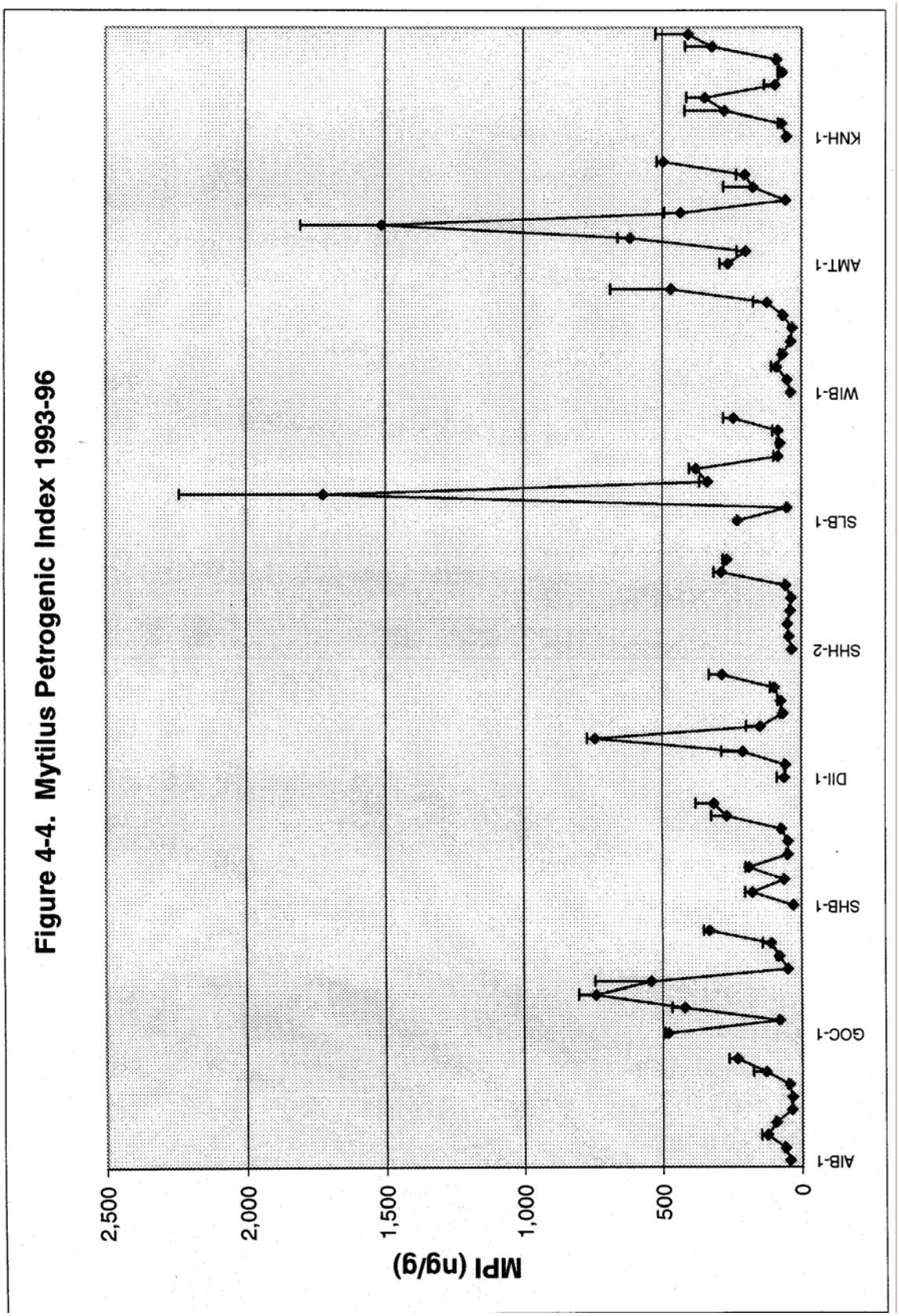


Figure 4-4. Mytilus Petrogenic Index 1993-96
MPI (ng/g)
2,500
2,000
1,500
1,000
500
0
AIB-1
GOC-1
SHB-1
DII-1
SHH-2
SLB-1
WIB-1
AMT-1
KNH-1

As shown by Figure 4-4, notable temporal changes and patterns occurred among the mussel samples collected at essentially every station. In examining this figure, it is important to note that the relative magnitude of the error bars associated with each triplicate measurement is very small compared to the overall change in *Mytilus* Petrogenic Index values observed for each station over time. Therefore, the observed trends are believed to reflect real changes in the field and not artifacts of the analytical method or collection procedure. As will be discussed below, the patterns observed at several of these stations can be correlated with spill events or clean-up activities that have occurred in Port Valdez or Prince William Sound since 1993.

4.3.2 Total PAH in *Mytilus* Tissues

Total PAH concentrations in mussel tissue samples throughout the measurement program in Prince William Sound ranged from 24 ng/g to 3,035 ng/g dry weight (Table 4-6). Generally, average TPAH concentrations in mussel tissue were low (<200 ng/g dry weight) in most samples analyzed (Figure 4-5) compared to TPAH concentrations reported for mussels in many parts of the world (e.g., Gosling 1992). However, TPAH concentrations noticeably higher than usual were observed at five sites on occasion. One site (Gold Creek) was categorized as reference, two were EVOS sites (Disk Island and Sleepy Bay), and two sites (Alyeska Marine Terminal and Knowles Head) were associated with tanker activities. Concentrations exceeded this value at all sites in March 1997, causing one to speculate on the possibility of procedural problems.

Lipid-corrected PAH concentrations in mussel tissue ranged from 290 ng/g lipid to 55,475 ng/g lipid. Generally, PAH concentrations in mussel tissue were low (<6,300 ng/g lipid) in most samples analyzed (Table 4-7 and Figure 4-6). Lipid-corrected PAH values exceeded 5,000 ng/g lipid on 27 occasions. Lipid-corrected PAH concentrations were noticeably higher than usual on occasion at five sites during the first eight surveys. The five noted sites included one reference sites, two EVOS sites, and two sites associated with tanker activities. As noted above, concentrations exceeded this value at all sites in March 1997.

4.3.2.1 Temporal

Short- or Long-term Trends

Highest concentrations of TPAH in tissues appeared to occur between March 1994 and March 1995 (Figure 4-5). Elevated levels were observed at Aialik Bay, Gold Creek, Disk Island, Sleepy Bay, and Alyeska Marine Terminal. The increased concentrations at the Alyeska Marine Terminal and Gold Creek sites appear to be a response to the *Eastern Lion* spill at the terminal in May 1994. It should be noted, however, that concentrations at both sites were rising in March 1994, prior to that spill, suggesting that some other source of hydrocarbons was available to mussels at that time.

Figure 4-5. Seasonal Variation in Total PAH in Mussel Tissues

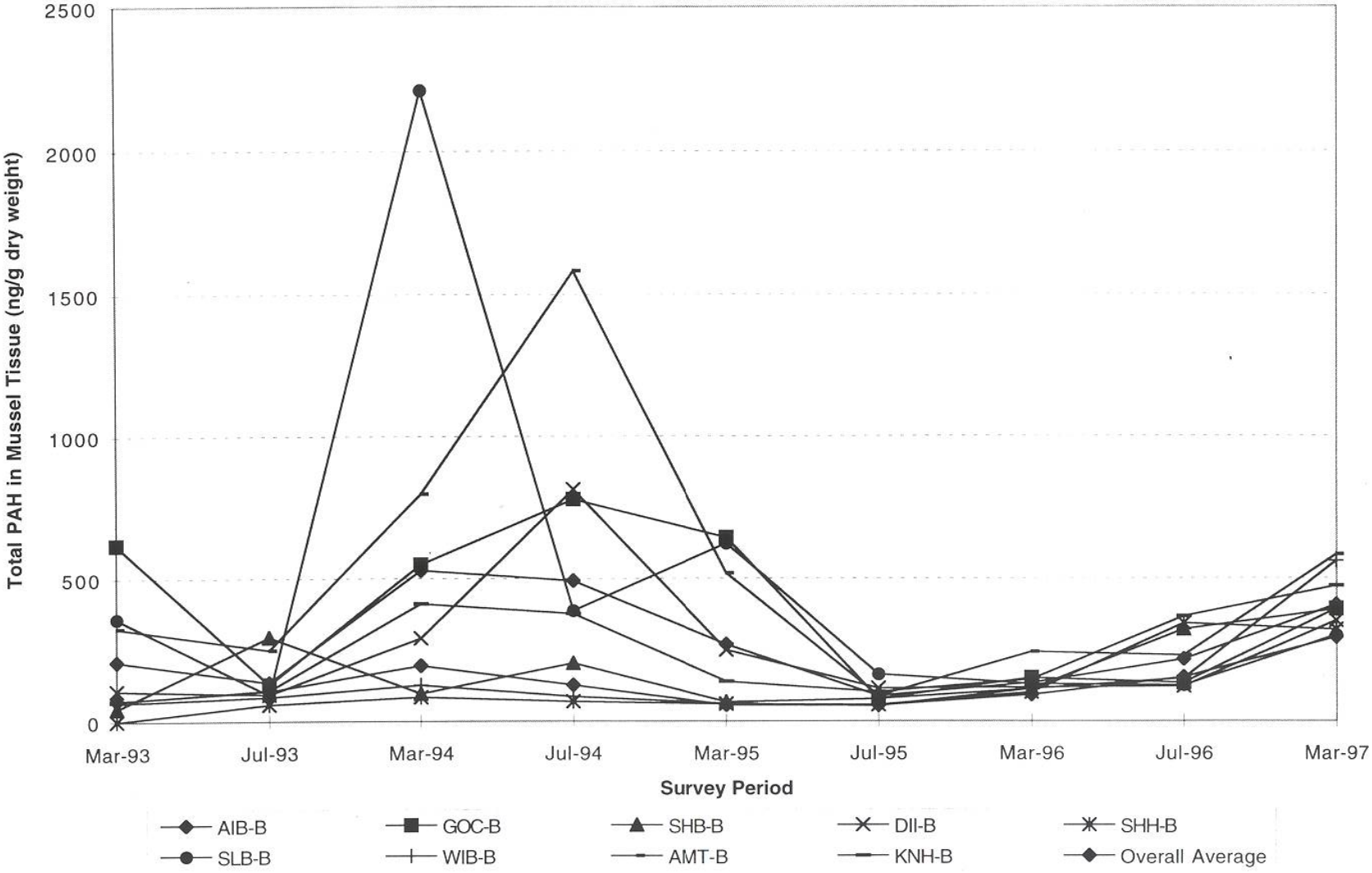
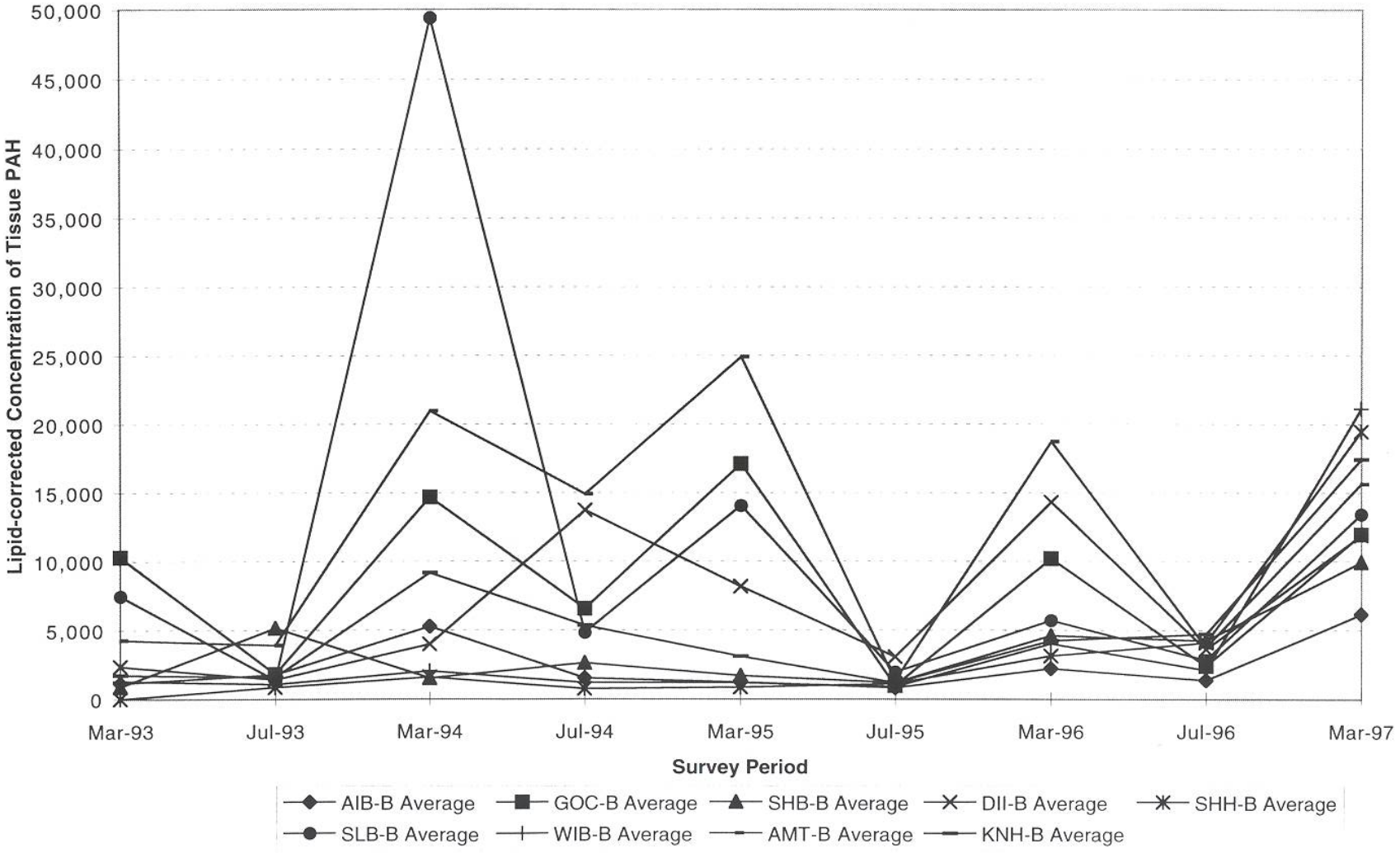


Table 4-7. Average Concentration of Lipid-corrected TPAH in Mussel Tissues in Samples from Specified Sites and Sampling Periods

Average Lipid-corrected TPAH in Mussel Tissue (ng/g dry lipid)

	SUMMER				WINTER					SUMMER	WINTER	Overall
Station/Survey Period	Jul-93	Jul-94	Jul-95	Jul-96	Mar-93	Mar-94	Mar-95	Mar-96	Mar-97	Average	Average	Average
										Category Averages		
Reference Sites										2526	6584	4780
Aialik Bay	1757	1504	825	1359	1136	5272	1197	2209	6143	1361	3191	2378
Gold Creek	1812	6531	969	2377	10311	14656	17081	10195	11944	2922	12837	8431
Sheep Bay	5134	2606	1193	4243	885	1518	1686	4593	9928	3294	3722	3532
Reference Average	2901	3547	996	2660	4111	7149	6654	5666	9338	2526	6584	4780
Std. Dev.	1934	2642	185	1462	5371	6767	9033	4100	2945	1026	5422	3214
Coeff. of Variation (%)	150	134	537	182	77	106	74	138	317	246	121	149
										Category Averages		
EVOS Sites										2791	9599	6510
Disk Island	1381	13708	3037	3589	2363	3964	8162	14305	11866	5429	8132	6931
Shuyak Harbor	842	722	1105	4089	NA	1566	830	3134	19438	1690	6242	3966
Sleepy Bay	1370	4802	1962	2815	7470	49403	14056	5694	13417	2737	18008	11221
Windy Bay	1070	1187	869	2104	1302	2006	1635	4019	21108	1308	6014	3922
EVOS Average	1094	2237	1312	3003	4386	17658	5507	4282	17987	1912	10088	6370
Std. Dev.	265	2233	575	1006	4361	27493	7415	1300	4045	740	6860	4201
Coeff. of Variation (%)	413	100	228	299	101	64	74	329	445	258	147	152
										Category Averages		
Tanker-related Sites										4586	12007	8709
Alyeska Marine Terminal	3897	14904	1333	3761	4272	20941	24883	18711	15630	5974	16887	12037
Knowles Head	1582	5316	1194	4703	1718	9160	3098	4230	17431	3199	7128	5382
Tanker Average	2739	10110	1263	4232	2995	15051	13990	11470	16531	4586	12007	8709
Std. Dev.	1636	6780	98	667	1806	8331	15404	10239	1274	1387	4880	3328
Coeff. of Variation (%)	167	149	1287	635	166	181	91	112	1298	331	246	262

Figure 4-6. Comparison of Temporal Patterns in Lipid-corrected PAH Concentrations



Mussels sampled at Disk Island in July 1994 also exhibited elevated concentrations of fresh ANS oil. As part of the EVOS cleanup, *Mytilus* beds near the sampling site had been rolled up a few days earlier to permit "cleansing" of the underlying sediment matrix. It is likely this cleanup operation inadvertently released a sheen of unweathered EVOS residues into the surrounding area, and that the mussel population sampled at Disk Island was exposed to this sheen.

Following that period, TPAH concentrations at all sites declined to low levels in July 1995 through March 1996. In fact, the background signals observed in the histogram plots for all mussel samples are strikingly similar (and low) in July 1995 and March 1996 (see Figure 3-2). This is also reflected in the uniformly low *Mytilus* Petrogenic Index values obtained in March and July 1995 (Figure 4-4). Levels remained fairly constant at most stations through July 1996, but concentrations were consistently elevated at all sites by March 1997. The mussels at Sleepy Bay and Disk Island exhibited PAH components consistent with ANS oil. The histogram patterns for the Alyeska Marine Terminal and Gold Creek are also consistent with relatively fresh ANS oil and probably reflect the spill at the Ballast Water Treatment Plant. However, the histogram patterns at Knowles Head and Sheep Bay are not consistent with ANS oil. The uniformity of this increase over the variety of sites sampled and the distances involved suggest that the general increase may represent a procedural problem or difference rather than a real increase in the background signals. As will be discussed in Section 4.7.2, elevated field and procedural blanks were noted for the March 1997 sampling period and analytical effort, and this could have contributed to some of the increases noted.

Seasonal Comparison

Generally, raw Total PAH concentrations were significantly lower in summer than in winter (Table 4-8; Figure 4-7; $p < 0.029$, one-way randomization t-test for dependent means) but the differences are not great. On average, this was true at reference and EVOS sites, but each category had exceptions as indicated by the generally large standard errors. Strong seasonal patterns were not observed at the sites associated with tanker activities. However, a greater number of the higher values occurred in winter than in summer. It is quite probable, however, that proximity to spill sources or disturbance events is more important than seasonal effects. Specifically, both low and high values were reported at Disk Island and Alyeska Marine Terminal during winters and summers (see the *Mytilus* Petrogenic Index plot, Figure 4-4) that cannot be easily rationalized without the benefit of specific spill events to explain the patterns.

In contrast, TPAH in tissues appeared to have a strong seasonal component when viewed on the basis of lipid content (Figure 4-6, Table 4-9). Values generally alternated between peaks during winter surveys and valleys during summer surveys; the pattern (values approximately 3 times higher in winter than in summer) was highly significant

($p < 0.001$, one-way randomization t-test for dependent means). When compared to temporal patterns in the uncorrected TPAH concentrations (Figure 4-5), it appears that the

Table 4-8. Seasonal relationships in TPAH among seasons and exposure categories.

	Mean TPAH ± SE (ng/g dry weight)	
Exposure Category	Summer	Winter
Reference Sites	204±61	251±135
EVOS Sites	175±48	318±158
Tanker-related Sites	386±212	370±173
Overall	231.6±47	307.2±73

lipid correction introduces an artificial element of variability. Furthermore, this correction seems to accentuate differences in TPAH concentrations in tissues among the three exposure categories (compare relationships in Tables 4-8 and 4-9).

4.3.2.2 Geographical

Concentrations of raw TPAH varied considerably among sites and between exposure categories (Tables 4-6 and 4-8). Concentrations at sites within the "so-called" reference and EVOS categories varied considerably spatially and temporally and differences were not significant in either summer ($p = 0.4$) or winter ($p = 0.5$). In contrast, the reference and tanker-related categories differed significantly in both summer ($p = 0.018$) and winter ($p = 0.019$) despite the high temporal and spatial variability among sites. The EVOS and tanker-related sites differed significantly during the summer ($p = 0.008$) but not in the winter ($p = 0.67$).

Likewise, concentrations of lipid-corrected TPAH varied considerably among exposure categories and among sites within the exposure categories (Tables 4-7 and 4-9). Reference and EVOS sites were not significantly different in summer or winter ($p = 0.85$ and 0.26, respectively). Reference and tanker-related sites were significantly different in both summer and winter ($p = 0.009$ and 0.074). As above, EVOS and tanker-related sites differed significantly in the summer ($p = 0.05$) but not in the winter ($p = 0.26$).

Mussel tissues at reference sites generally exhibited low concentrations of PAHs except at Gold Creek. At that site, PAH concentrations were substantially higher than average in 4 of 9 surveys and the pattern generally correlated closely with those observed in mussel tissues from the Alyeska Marine Terminal (Figure 4-5).

Figure 4-7. Comparison of PAH in Mussel Tissue by Season

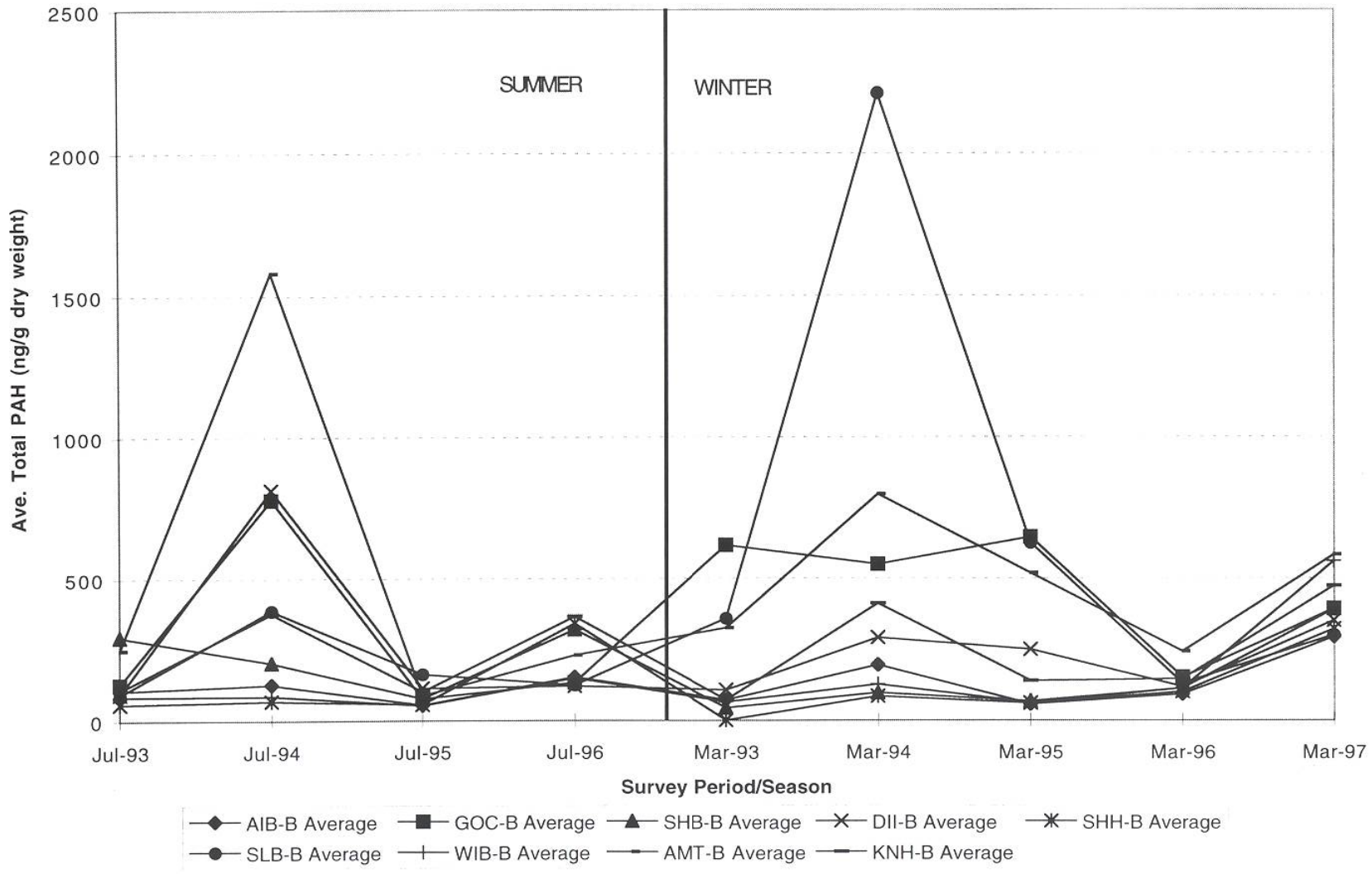


Table 4-9. Seasonal relationships in lipid-corrected TPAH among seasons and exposure categories.

	Mean TPAH ± SE (ng/g lipids)	
Exposure Category	Summer	Winter
Reference Sites	2,526±2,054	6,584±1,221
EVOS Sites	2,791±1,397	9,577±2,811
Tanker-related Sites	4,583±2,240	12,007±7,411
Overall	3,102±588	9,119±1,940

Two of the EVOS sites exhibited concentrations that were substantially higher than the overall average. Sleepy Bay stood apart in 5 out of 9 surveys and Disk Island stood apart in 3 out of 9 surveys. In contrast, the far-field EVOS sites (Windy Bay and Shuyak Harbor) were characterized by low concentrations in most surveys (Figure 4-5, Table 4-6).

Both sites associated with tanker activities exhibited TPAH concentrations noticeably higher than average during at least four surveys. The Alyeska Marine Terminal exhibited the strongest signal and was noticeably higher in 5 of 9 surveys. Knowles Head, the anchorage in central PWS, exhibited a substantially weaker signal and was noticeably higher in 4 of 9 surveys. Typically all five of these sites exhibited the high concentrations during consecutive surveys from March 1993 through March 1995, when known sources of ANS-derived PAH were introduced into PWS.

Patterns were similar when viewed from the perspective of lipid-corrected TPAH. However, most of the sites with high values exhibited dramatic seasonal changes in lipid-corrected PAH concentrations. Many of these sites exhibited the high concentrations during consecutive winter surveys from March 1993 through March 1997. Mussel tissues at reference sites generally exhibited low concentrations of lipid-corrected PAHs (<5,000 ng/g lipid) except at Gold Creek. At that site, lipid-corrected PAH concentrations were

markedly higher in 6 of 9 surveys and again generally correlated closely with TPAH in tissues at the Alyeska Marine Terminal. Highest values were recorded in the one-year period from March 1994 through March 1995.

Tissue concentrations of TPAH were substantially above average at all sites (17 to 27 percent) in March 1997. As indicated above, this suggests that the increases observed at some of these sites were the result of procedural differences or problems rather than reflections of real changes in Prince William Sound. It is highly unlikely that a change of the magnitude observed in the types of hydrocarbons "sampled" by mussels occurred over the entire geographic region sampled.

4.4 Evaluation of Mytilus *Biological Data*

4.4.1 Morphometrics

In terms of mensural characteristics of the mussel samples, the populations collected appear generally suitable. The three measures of size were length, volume, and dry tissue weight. Average length exceeded 25 mm in all samples (Figure 4-8), a length assuring reproductive maturity (Arkhipova, 1992; Houghton et al., 1992). The animals from Knowles Head, with sample populations averaging from 25.4 to 32.9 mm and averaging 28.9 mm, were smallest in every survey. The animals from the Alyeska Marine Terminal, Gold Creek, Disk Island, and Shuyak Harbor, with average length ranging from about 35 mm to nearly 45 mm, were generally among the larger specimens (Figure 4-8).

Average length varied substantially among stations, but lengths were relatively consistent among surveys at each specific site (Figure 4-8). However, length generally fluctuated seasonally at most of the stations so that average size was larger in summer and smaller in winter. No substantive trends were observed in shell length over the period surveyed.

In view of relationships between mussel size and rates of physiological and metabolic processes, it is likely that hydrocarbon concentrations in tissues are influenced by mussel size. To account for this, the NOAA Mussel Watch program sets a specific size guideline for sample collection (O'Connor et al. 1994). While TPAH concentrations did not exhibit correlations with size at specific sites over the course of this monitoring program (Aialik Bay - $r = 0.38$; Shuyak Harbor - $r = 0.27$; Windy Bay - $r = 0.58$), it would be prudent to establish a narrower range of acceptable shell sizes for the mussel collections.

Length and volume are strongly correlated (Figure 4-9; $r = 0.93$; $p << 0.0001$). Thus the patterns and relationships noted above for length are valid for volume, making volume a superfluous measurement.

The most important use of size data in this program is to confirm that the populations of sample animals are within an acceptable size range. For this purpose, a single measurement, e.g., length, will suffice. Based on the method described by KLI for

Figure 4-8. Temporal Patterns in Shell Length of Sampled Mussels

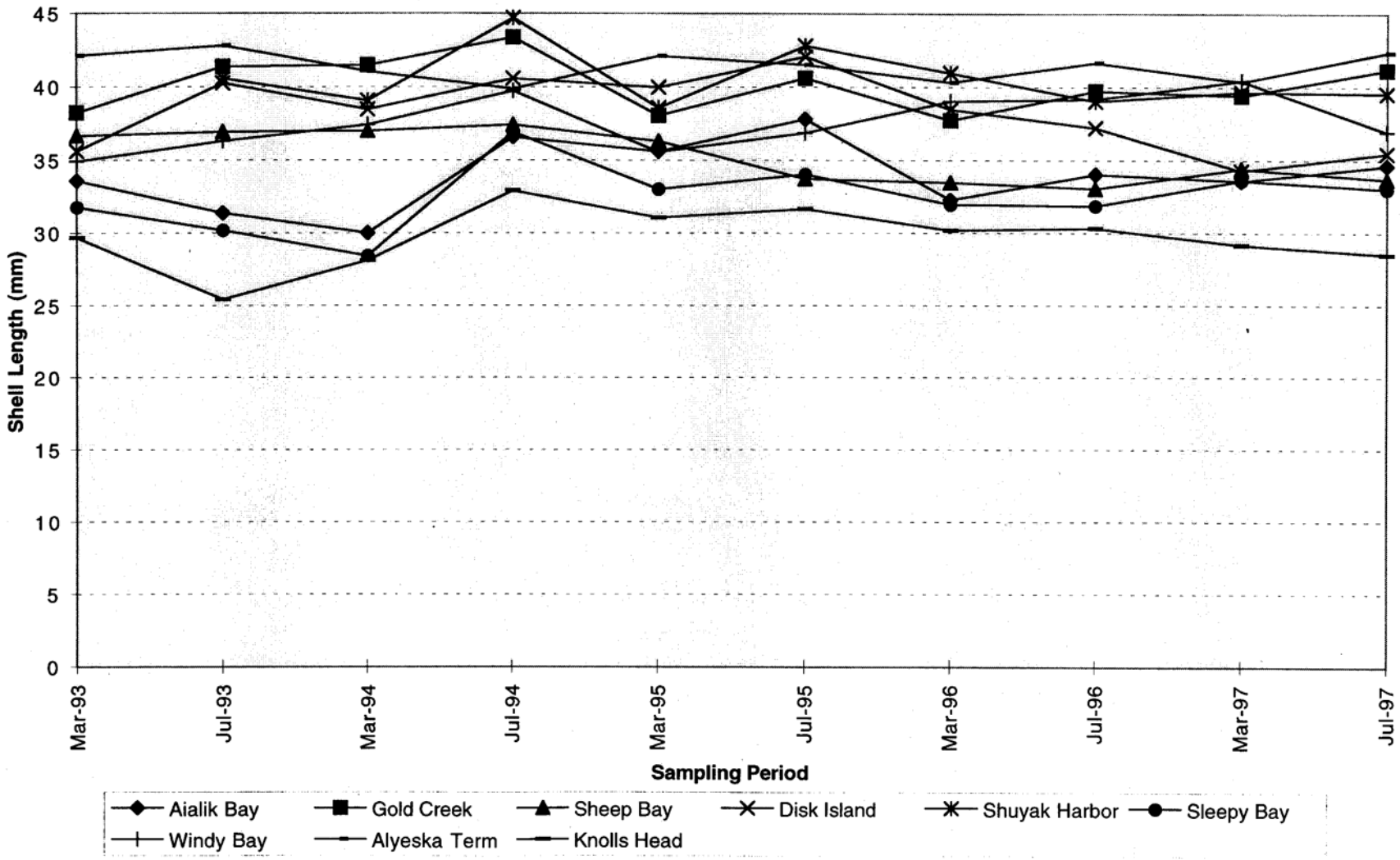
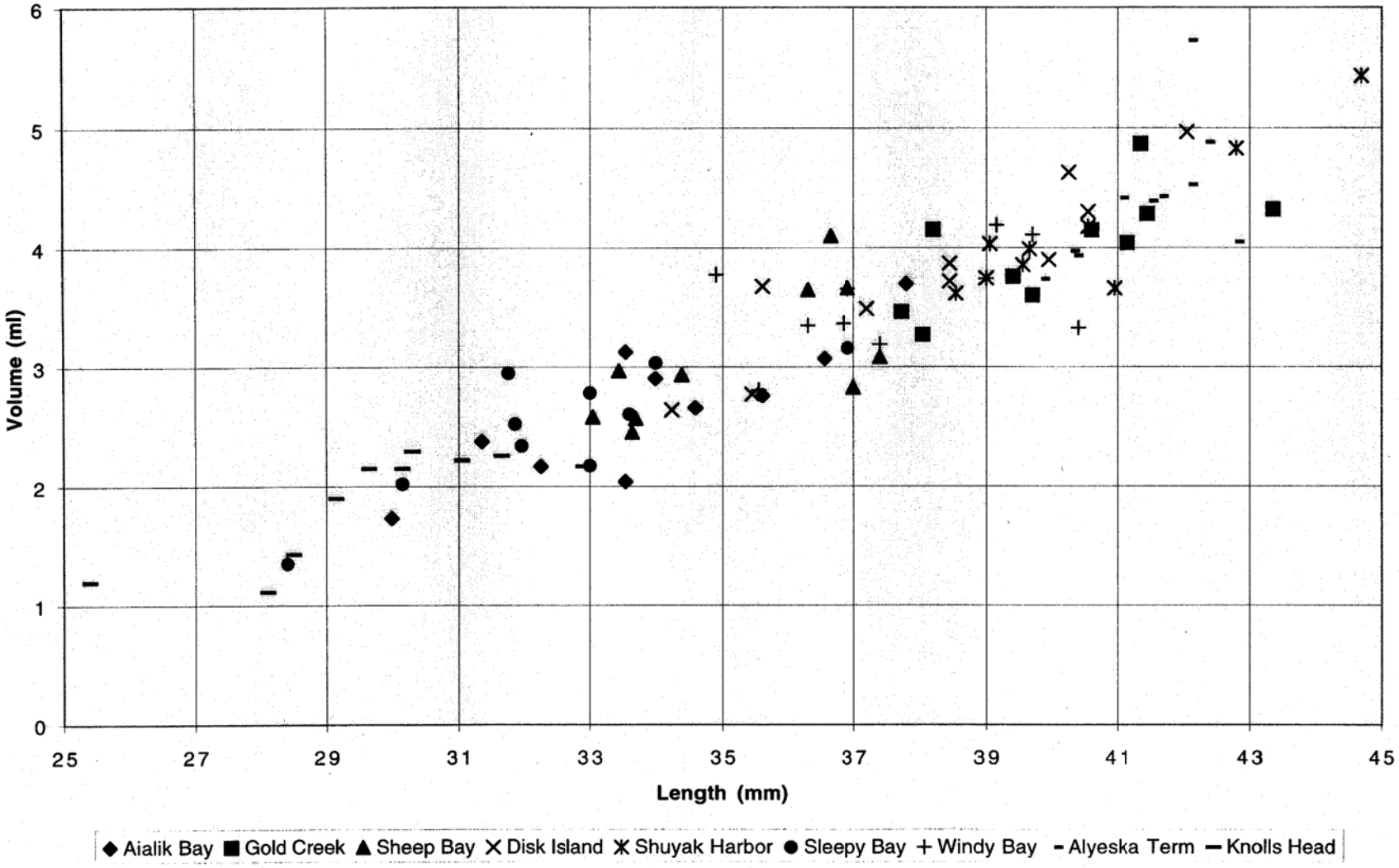


Figure 4-9. Length-Volume Relationship for Mussels



measuring volume, it is likely that length measurements are more accurate and precise. The remaining measurements should be omitted from the lab work up unless other justifications for their continuation are identified.

4.4.2 Condition and Tissue Development

Tissue development was evaluated using gonadal and somatic (non-gonadal) tissue weight, and pooled volumes for gonadal and somatic tissues from the samples at each site. However, no rationale other than consistency with the NOAA Mussel Watch program has been identified for collecting these kinds of data. The purpose for measuring tissue weights is questionable in view of the objectives of this program. The sensitivity of these measurements is poor, especially considering the low frequency of sampling. Far better techniques are available to measure stress or exposure if that were the objective. However, the purpose of this program is to monitor petroleum hydrocarbons in sediment and tissues rather than evaluate the condition of the populations. We therefore recommend that collection of these ancillary types of data be terminated and the funds be diverted to expand more useful elements of the program.

4.4.3 Lipid Measurements and Relationships

Because hydrocarbons are lipophilic, percent lipids have been measured to permit standardization of the Total PAH measurements on a lipid basis. In fact, lipid-corrected data appear to confuse the picture. Lipid content ranged between 2.65 and 13.1 percent of dry tissue weight in summer and 0.67 and 8.4 percent in winter. Generally, lipid content alternated between high percentages in summer and low percentages in winter (Figure 4-10). This alteration in lipid content is a response to the balancing of nutrition availability, metabolic rates, and reproductive activities (Thompson, 1984; Kimball, 1993; Dare, 1975). The levels of lipids observed are within the range generally reported for mussels (e.g., Gosling 1992).

The only potentially useful comparison for TPAH and the various characteristics measured for mussels (i.e., length, volume, and gonadal, somatic, and tissue weight) is with lipids. TPAH and lipids are measured for the same composite sample, whereas the other characteristics are either measured for pooled samples or individual specimens that were not analyzed chemically. Thus, the relationship between the chemical and the biological measurements is not valid. Consequently, none of these variables except lipid content is useful either analytically or predictively for evaluating TPAH in mussel tissues.

The relationship between lipid content and TPAH concentrations in mussel tissue from sites with minimal exposure to hydrocarbons (Figure 4-11) was opposite to that predicted by the often-cited positive correlation between lipids and lipophilic (hydrophobic) organic compounds. Moreover, the relationship between lipids and TPAH appears weak, and TPAH varied more widely when viewed on the basis of lipids (Figure 4-6). In contrast, body burdens of TPAH in mussels at these lightly exposed sites appears to be relatively stable (Figure 4-5). Sericano et al. (1990) and Widdows et al. (1990) have both reported that hydrophobic contaminants show little correlation with lipids in mussels. The

Figure 4-10. Temporal Comparison of Percent Lipids in Mussel Tissue

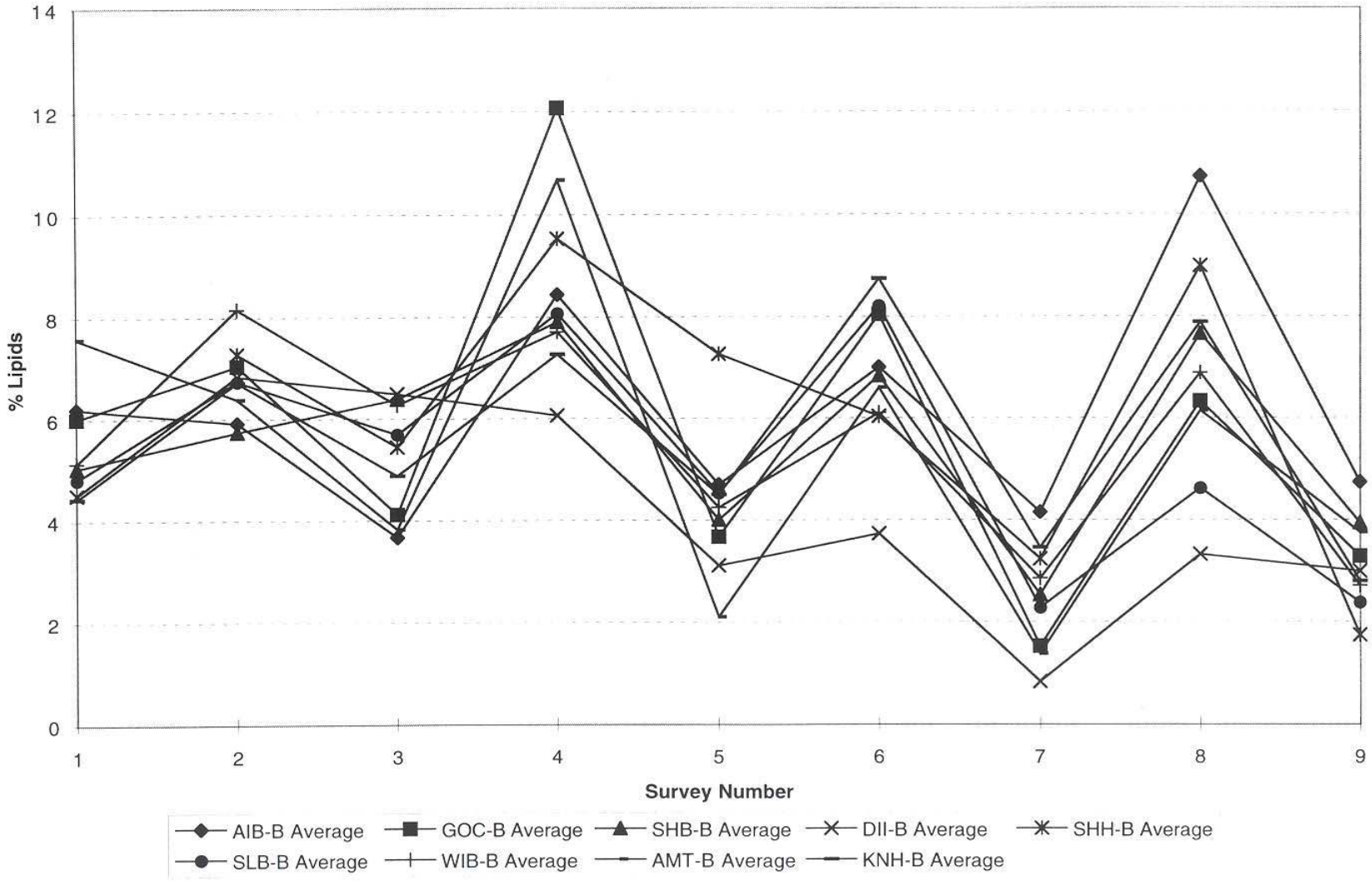
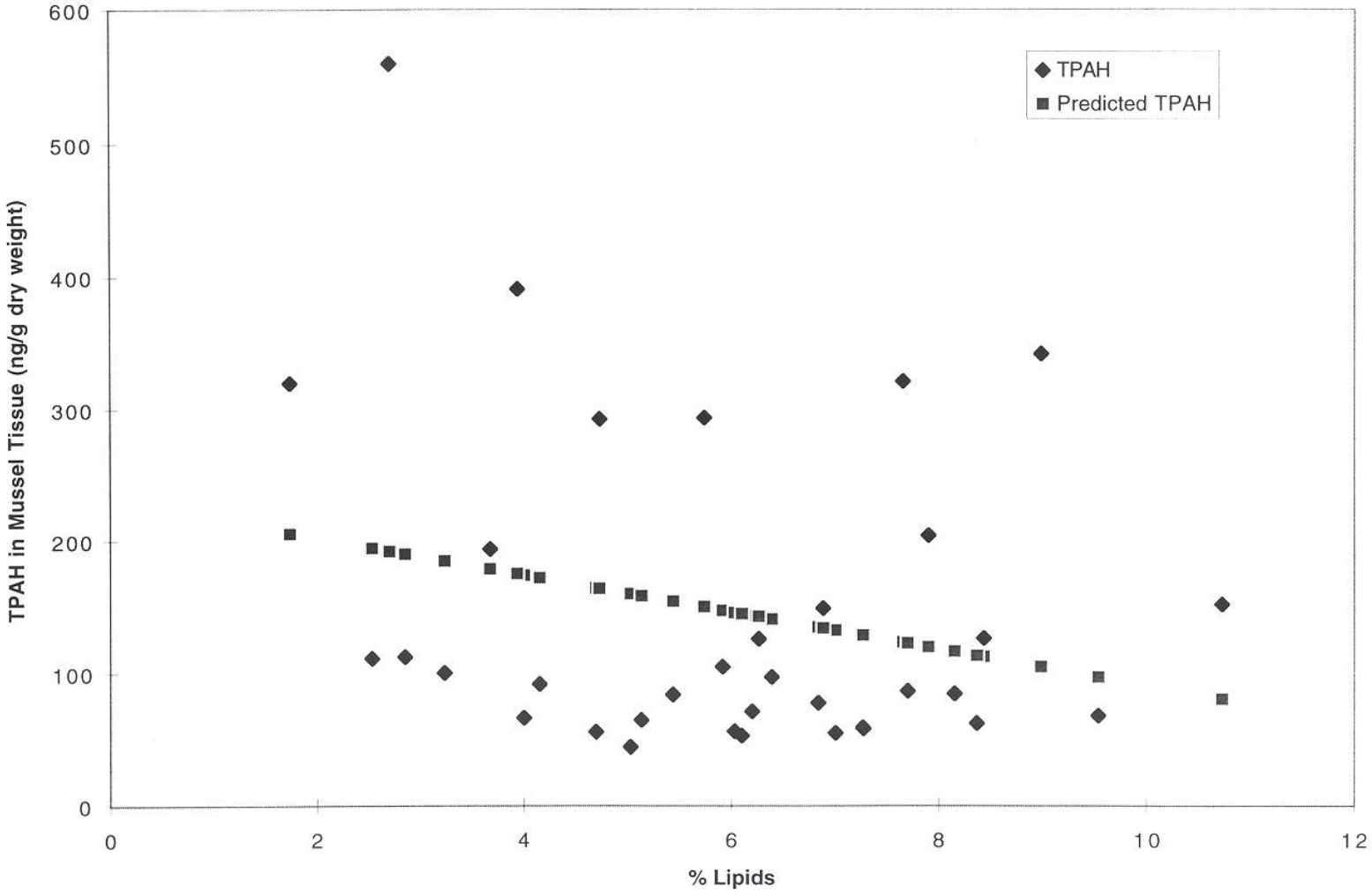


Figure 4-11. Relationship Between TPAH and % Lipids for Mussels from Aialik Bay, Sheep Bay, Shuyak Harbor, and Windy Bay



relationship between lipid content and both gonad and somatic weight is weak (e.g., Figure 4-12), which is not surprising in view of the fact that lipid content is generally much higher in tissues of the mantle and digestive gland than in gonads in mussels (Gosling 1992). These facts seem to provide compelling arguments in favor of a recommendation to discontinue lipid analysis of mussel tissues in the LTEMP.

4.4.4 Hydrocarbon Signatures and Minimum Detection Limit Issues

Assessing hydrocarbon contamination in mussel tissue using Total PAH is a rather crude approach, especially considering that most of the values for TPAH are somewhat below the summed MDL for the target constituents. In fact, summed MDLs were approximated (within 20 ng/g dry weight TPAH) or exceeded in only 15 of the 80 sample sets analyzed (Table 4-10). These 15 hits were distributed among Gold Creek, Sleepy Bay, Disk Island, and Alyeska Marine Terminal in March 1993; March and July 1994, March 1995, and March 1997. The only summer survey to exhibit hits was July 1994. Six of these hits (and all of the hits occurring in summer) can be related to known events wherein petroleum hydrocarbons were released into the water column. It is likely, for example, that the July 1994 hits at Alyeska Marine Terminal and Gold Creek were a response to the *Eastern Lion* spill at the terminal in May 1994. Moreover, it is likely that the March 1997 hits at the terminal and Gold Creek were a response to the spill at the Ballast Water Treatment Plant in January 1997. The July 1994 hits at Disk Island are probably responses to EVOS-related cleanup efforts in adjacent areas (i.e., removal of an oil-saturated mussel bed at Disk Island). In contrast, the hits in March 1993, 1994, 1995, and 1997 at Sleepy Bay and Disk Island, two sites in Prince William Sound that were heavily contaminated by EVOS, may have resulted from disturbance of buried EVOS oil residues or contaminated intertidal sediments by violent winter storms.

In an effort to examine the data in more detail for evidence of real hits and insight into the nature of the exposures, 32 selected parent compounds and C-1 through C-4 homologues were examined at each site to determine where hydrocarbon signals exceeded the MDLs, despite the overall indication from the sum of the MDLs (Table 4-10) that hydrocarbons were below detection. The sampling dates on which each of these selected compounds exceeded its survey-specific MDL are listed in Table 4-11. Only ten constituents occurred above their individual MDLs at every site at some time during the program. These included anthracene, biphenyl, C-1 naphthalene, C-3 fluorene, C-3 naphthalene, C-4 naphthalene, dibenzothiophene, fluoranthene, naphthalene, and phenanthrene. Several other compounds routinely occurred at levels above their individual MDLs at seven or eight sites, including benzo(a)anthracene, C-1 fluorene, C-2 chrysene, C-2 fluorene, and C-2 naphthalene. Naphthalene was the only constituent that occurred above its MDL at every site during every survey. Also, phenanthrene exceeded its MDL in most cases.

A hydrocarbon exposure index was calculated by multiplying the number of constituents at a site that exceeded their individual MDLs by the frequency with which exceedance occurred (see Table 4-11). Based on this approach, the sites at which exposure to hydrocarbons was most apparent in mussel tissues were Sleepy Bay, Gold Creek, Alyeska Marine Terminal, and Disk Island. The sites at which mussels were least exposed to

Figure 4-12. Correlation Between Gonad Weight and Lipid Content

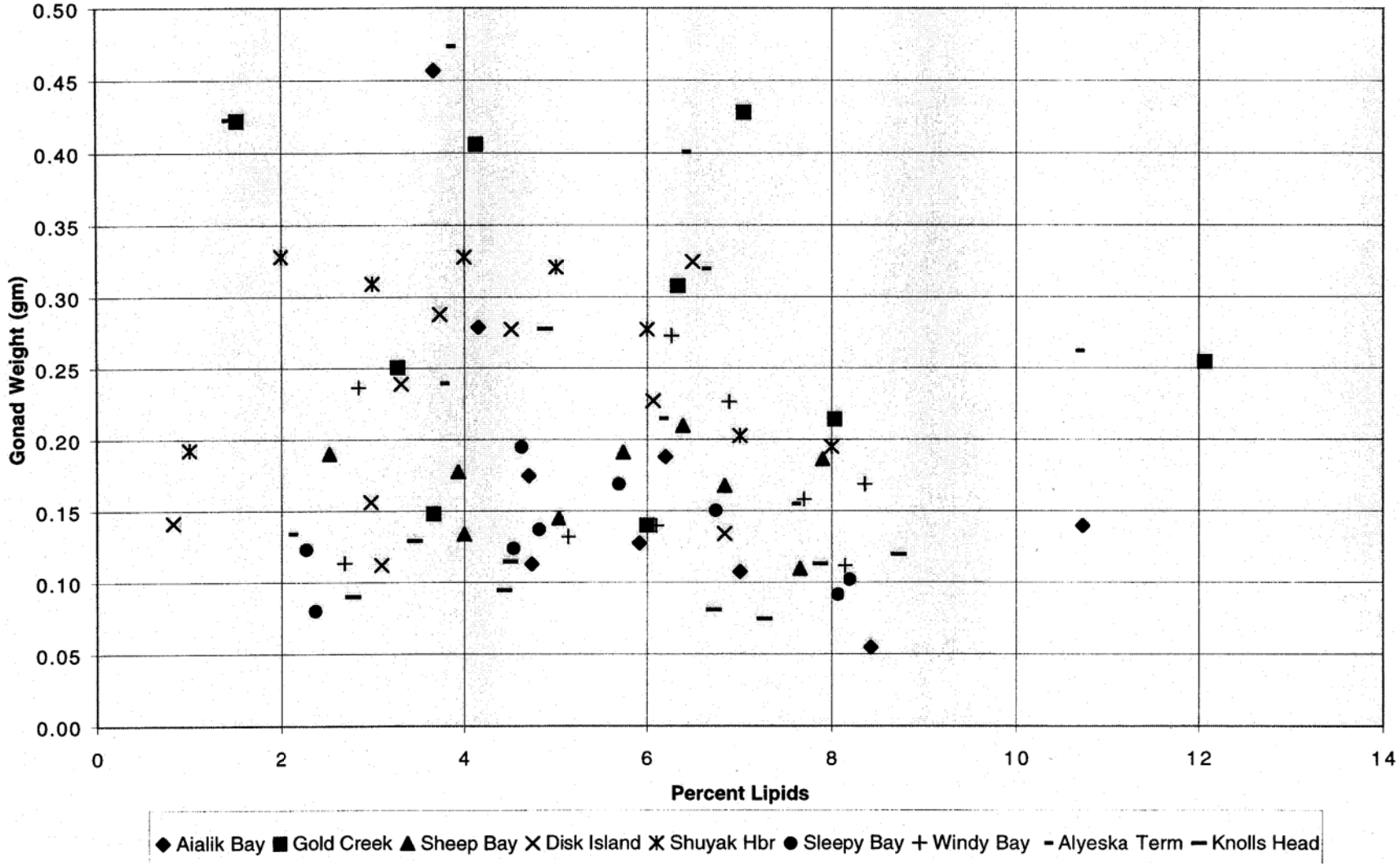


Table 4-10. Difference between measured TPAH concentration in mussel tissue for the specified station from the specified survey period and the estimated sum of the Method Detection Limits for the target PAHs for that sample. Bold values indicate that the measured value was higher than the MDL for TPAH and negative values indicate cases where the estimated MDL exceeded the measured value.

	Survey Period								
Station ID	Mar-93	Jul-93	Mar-94	Jul-94	Mar-95	Jul-95	Mar-96	Jul-96	Mar-97
Reference Sites									
Aialik Bay	-249	-557	-433	-202	-246	-229	-202	-163	-185
Gold Creek	**393**	-685	**-12**	**497**	**37**	-292	-185	-210	**42**
Sheep Bay	-259	-984	-261	-117	-635	-332	-240	-177	-340
EVOS Sites									
Disk Island	-340	-515	**-14**	**562**	-96	-420	-178	-210	**-17**
Shuyak Harbor	NS	-432	-202	-110	-255	-289	-231	-209	-268
Sleepy Bay	**124**	-495	**1435**	**88**	**148**	-408	-152	-328	-225
Windy Bay	-256	-406	-226	-167	-276	-232	-243	-290	-321
Tanker Sites									
Alyeska Marine Terminal	**39**	-445	**372**	**1322**	-115	-360	-57	-69	**236**
Knowles Head	-66	-97	-373	-365	-123	-84	-119	-349	-459
Average	-77	-513	32	168	-174	-294	-179	-223	-171
SE	95	84	204	189	79	37	22	31	77

Table 4-11. Sampling periods in which selected PAH constituents exceeded site-specific MDLs in mussel tissues.

Constituents	Survey Periods at Sampling Sites									Number of Sites
	AIB-B	GOC-B	SHB-B	DII-B	SHH-B	SLB-B	WIB-B	AMT-B	KNH-B	
Anthracene	3-7/94; 7/95	3/93; 3-7/94; 3/95; 3/96	3/94; 3-7/96	3/94; 3-7/95; 3/96	3/94; 3/96	3/94; 3-7/95; 3/96	3/94; 7/95; 3/96	3-7/94; 7/95; 3/96	7/95; 3/96	9
Benzo (a) anthracene	3/96	3/93; 3/94; 3/96		3/95	3/96	3/93; 3/94; 3/95; 3-7/96		3/93; 3/96	3/96	7
Benzo (a) pyrene			7/96			3/95				2
Benzo (a) fluoranthene		3/93; 3/95; 3/97		3/95; 3/97		3/94; 3/95				3
Benzo (e) pyrene	7/93; 7/94	3/93; 3/94; 3/95	3-7/94	3-7/93; 3-7/94; 3/95		3-7/93; 3-7/94; 3-7/95; 3/96		3/93; 3-7/94; 3/95; 3/96		6
Benzo (k) fluoranthene		3/93		3/95		3/94; 3/95				3
Biphenyl	3-7/94; 3/95; 3/97	7/94; 3/96; 3/97	7/94; 3/97	7/94; 3/95; 3/97	3-7/94; 7/96; 3/97	3-7/94; 3/97	3-7/94; 3/95; 3/96; 3/97	3-7/94; 3/97	3-7/94; 3/96; 3/97	9
C1-chrysene		3/93		3-7/94		3/93; 3/94; 3/95		3/93; 3-7/94; 3/97		4
C1-dibenzothiophene		3/93; 3-7/94; 3/95		3-7/94		3-7/94; 3/95		3/93; 3-7/94; 3/95		4
C1-fluoranthene/pyrene		3/93; 3-7/94	3/94	3-7/94	3/94	3/93; 3-7/94; 3-7/95		3/94; 3/97		6
C1-fluorene		7/94	7/94; 7/96	7/94	7/96	7/94		7/94	7/94	7
C1-naphthalene	3/93; 3-7/94; 7/95; 3-7/96	3/93; 3-7/94; 3/96	3-7/93; 3-7/94; 7/95; 3-7/96	3-7/94; 3-7/96	3-7/94; 3-7/96	3/93; 3-7/94; 7/95; 3/96	3/93; 3-7/94; 3-7/96	3-7/93; 3-7/94; 3/96	3/93; 3-7/94; 7/95; 3-7/96	9
C1-phenanthrene		3/93; 7/94				3/94		3/93; 7/94		3
C-2 chrysene	3/97	3/93	3/97	3-7/94; 3/97	3/97	3/93; 3/94; 3/95		3/93; 3-7/94; 3/97		7
C-2 dibenzothiophene		3/93; 3-7/94; 3/95; 3/97		3-7/94; 3/95; 3/97		3/93; 3-7/94; 3/95		3/93; 3-7/94; 3/95; 3/97		4
C-2 fluorene	7/96	3/93; 7/94; 3/97	7/94; 7/96	7/94; 3/97	7/96	3-7/94		7/94; 7/96; 3/97	7/94; 7/96	8
C-2 naphthalene		3/93; 7/94; 3/95; 3/97	7/94; 7/96	7/94; 3/97		7/94; 3/97	3/94; 3/97	7/93;3-7/94; 3/97	7/94; 7/96; 3/97	7
C-2 phenanthrene		3/93; 3-7/94; 3/95; 3/97		3-7/94		3/93; 3-7/94; 3/95		3/93; 3-7/94; 3/95; 3-7/96; 3/97		4
C-3 chrysene										0

Table 4-11. Sampling periods in which selected PAH constituents exceeded site-specific MDLs in mussel tissues.

Constituents	Survey Periods at Sampling Sites									Number of Sites
	AIB-B	GOC-B	SHB-B	DII-B	SHH-B	SLB-B	WIB-B	AMT-B	KNH-B	
C-3 dibenzothiophene		3/93; 3-7/94; 3/95; 3/97		3-7/94; 3/95; 3/97		3/93; 3-7/94; 3/95		3/93; 3-7/94; 3/95; 3-7/96; 3/97		4
C-3 fluorene	7/96; 3/97	3/93; 7/94; 3/95; 3/97	7/96; 3/97	7/94; 3/97	7/96; 3/97	3/94	7/96; 3/97	3/93; 7/94; 3/95; 7/96; 3/97	7/94; 7/96; 3/97	9
C-3 naphthalene	3/97	3/93; 7/94; 3/95; 3/97	7/94; 7/96; 3/97	7/94; 3/97	3/97	7/94; 3/97	3/94; 3/97	7/93; 3-7/94; 3/97	7/94; 7/96	9
C-3 phenanthrene		3/93; 3-7/94; 3/95; 3/97		3-7/94; 3/95; 3/97		3/93; 3-7/94; 3/95		3/93; 3-7/94; 3/95; 3-7/96; 3/97		4
C-4 chrysene										0
C-4 naphthalene	3/97	3/93; 7/94	7/94; 3/97	7/94	3/97	3-7/94	3/97	3-7/94; 3/97	3/97	9
C-4 phenanthrene		3/93; 3-7/94; 3/95; 3/97		3-7/94; 3/97		3/93; 3-7/94; 3/95		3/93; 3-7/94; 3/95; 3-7/96; 3/97		4
Chrysene		3/93; 3/94; 3/95; 3/96; 3/97		3/94; 3/95; 3/97		3/93; 3/94; 3/95; 3/96		3/93; 3/94; 3/96; 3/97		4
Dibenzothiophene	3-7/94; 7/95	3/93; 3-7/94; 7/95	3/94; 7/95	3-7/94; 7/95	3-7/94; 7/95; 3/96	3/94; 7/95; 3/96	3-7/94	3-7/94; 7/95	3/94	9
Fluoranthene	3-7/94; 3/96	3/93; 3-7/94; 3-7/95; 3/96; 3/97	3/94; 3/96	3/94; 3/95; 3/96; 3/97	3-7/94	3/93; 3/94; 3-7/95; 3/96	7/93; 3-7/94	3/93; 3-7/94; 3-7/95; 3/96	3/96	9
Fluorene	3-7/94; 3/96		3/96		3/96; 3/97	3/94; 3/96			3/94; 3/96	5
Naphthalene	All surveys	All surveys	All surveys	All surveys	All surveys	All surveys	All surveys	All surveys	All surveys	9
Phenanthrene	All but 3/97	All but 3/97	All but 3/97	3-7/93; 3-7/94; 3-7/95; 3/96	All but 3/97	All but 3/97	All but 3/97	All but 3/97	All but 3/97	9
No. of Constituents	15	28	18	27	16	30	11	26	15	
Sum of Occurrences	48	109	52	82	44	103	42	118	46	
Hydrocarbon Exposure Index	720	3052	936	2214	704	3090	462	3068	690	
Pyrogenic Index*	0.21	0.17	0.12	0.18	0.16	0.25	0.12	0.12	0.15	

* Pyrogenic Index = Sum of no. of occurrences for benzo (a) anthracene, benzo (a) pyrene, benzo (e) pyrene, benzo (b) fluoranthene, benzo (k) fluoranthene, biphenyl, chrysene, dibenzo (ah) anthracene, and fluorene divided by sum of no. of occurrences for all indicated constituents

hydrocarbons were Windy Bay, Knowles Head, Shuyak Harbor, Aialik Bay, and Sheep Bay (Table 4-11). Ordering the sites according to increasing value of the index provides a useful comparison among the sites of the relative frequency of contamination (Table 4-12). Based on this index, the sites appear to fall into two distinct groups. The hydrocarbon exposure index ranges from 462 to 936 for the less exposed group and from 2,214 to 3,090 for the more exposed group. The less exposed group includes three sites that were originally selected to represent exposure to EVOS contamination or tanker activities. The more exposed group includes one site originally selected as a reference for the Alyeska Terminal (Gold Creek) as well as the terminal and two EVOS sites that have exhibited persistent hydrocarbon releases since the completion of the 1989-90 shoreline treatment programs.

The importance of pyrogenic compounds was evaluated with an index assessing the relative frequency of occurrence of pyrogenic constituents at each site. The pyrogenic index ranges from 0.12 to 0.21 for the less exposed group and from 0.12 to 0.25 for the more exposed group. Generally, concentrations of pyrogenic compounds in mussel tissues did not appear to correlate with higher hydrocarbon exposure indices or average TPAH concentrations at the sites (Table 4-11). That is, the pyrogenic compounds were rather ubiquitous throughout the study region.

Constituents were more likely to be observed at levels above their MDLs in winter surveys than in summer (Table 4-11). Sixty-three percent of these cases occurred in winter vs. 37 percent in summer, i.e., it was nearly twice as likely that constituents were detected above MDLs in winter as in summer. It is not clear whether this indicates a greater likelihood of detecting exposure or that exposure occurs more frequently in winter. For the constituents examined, MDLs were exceeded in approximately 30 percent of the analyses.

It is important to emphasize that, generally, the concentrations of hydrocarbons observed in mussel tissue in this program are relatively low. As pointed out above, nearly 70 percent of the constituents were measured below their pertinent MDLs. By way of comparison, concentrations of TPAHs measured in mussel tissues in 1991 for some EVOS sites in Prince William Sound ranged between 360 (Eshamy Bay) and 20,000 ng/g dry weight (Smith Island). A mussel sample collected near the cannery in Seward had a TPAH concentration of 6,200 ng/g dry weight (Houghton et al. 1992). The highest average TPAH concentration observed in this program was 2,209 ng/g dry weight at Sleepy Bay, and only one other measurement exceeded 1,000 ng/g dry weight.

Moreover, it is quite likely that the higher values at several sites reflect a response to local activities rather than widespread contamination. The two EVOS sites in Prince William Sound are located in areas with persistent release of EVOS oil sequestered in the sediments and where cleanup activities disturbed the sediments repeatedly until at least 1994. The importance of pyrogenic compounds relative to petrogenic compounds is a

Table 4-12. Number of surveys in which selected PAH constituents exceeded site-specific MDLs in mussel tissues.

Constituents	Sampling Sites									No. of Sites	Total Occurrences
	WIB-B	KNH-B	SHH-B	AIB-B	SHB-B	DII-B	GOC-B	AMT-B	SLB-B		
Anthracene	3	2	4	3	3	4	5	4	2	9	30
Benzo (a) anthracene		1	1	1		1	3	2	5	7	14
Benzo (a) pyrene					1				1	2	2
Benzo (a) fluoranthene						2	3		2	3	7
Benzo (e) pyrene				2	2	5	3	5	7	6	24
Benzo (k) fluoranthene						1	1		2	3	4
Biphenyl	5	4	4	4	2	3	3	3	3	9	31
C1-chrysene						2	1	4	3	4	10
C1-dibenzothiophene						2	4	4	3	4	13
C1-Fluoranthene/pyrene			1		1	2	3	2	5	6	14
C1-fluorene		1	1		2	1	1	1	1	7	8
C1-naphthalene	5	6	4	6	7	4	4	5	5	9	46
C1-phenanthrene							2	2	1	3	5
C-2 chrysene			1	1	1	3	1	4	3	7	14
C-2 dibenzothiophene						4	5	5	4	4	18
C-2 fluorene		2	1	1	2	2	3	3	2	8	16
C-2 naphthalene	2	3			2	2	4	4	2	7	19
C-2 phenanthrene						2	5	7	4	4	18
C-3 chrysene										0	0
C-3 dibenzothiophene						4	5	7	4	4	20
C-3 fluorene	2	3	2	2	2	2	4	5	1	9	23
C-3 naphthalene	2	2	1	1	3	2	4	4	2	9	21
C-3 phenanthrene						4	5	7	4	4	20
C-4 chrysene										0	0
C-4 naphthalene	1	1	1	1	2	1	2	3	2	9	14
C-4 phenanthrene						3	5	7	4	4	19
Chrysene						3	5	4	4	4	16
Dibenzothiophene	2	1	4	3	2	3	4	3	3	9	25
Fluoranthene	3	1	2	3	2	4	7	6	5	9	33
Fluorene		2	2	3	1				2	5	10

Table 4-12. Number of surveys in which selected PAH constituents exceeded site-specific MDLs in mussel tissues.

Constituents	Sampling Sites									No. of Sites	Total Occurrences
	WIB-B	KNH-B	SHH-B	AIB-B	SHB-B	DII-B	GOC-B	AMT-B	SLB-B		
Naphthalene	9	9	8	9	9	9	9	9	9	9	80
Phenanthrene	8	8	7	8	8	7	8	8	8	9	70
No. of Constituents	11	15	16	15	18	27	28	26	30		
Sum of Occurrences	42	46	44	48	52	82	109	118	103		
Hydrocarbon Exposure Index	462	690	704	720	936	2214	3052	3068	3090		
Pyrogenic Index*	0.12	0.15	0.16	0.21	0.12	0.18	0.17	0.12	0.25		
Average TPAH (ng/g dry tissue)	144	243	136	127	178	250	386	512	487		

* Sum of no. of occurrences for benzo (a) anthracene, benzo (a) pyrene, benzo (e) pyrene, benzo (b) fluoranthene, benzo (k) fluoranthene, biphenyl, chrysene, dibenzo (ah) anthracene, and fluorene divided by sum of no. of occurrences for all indicated constituents

reflection of how low the concentrations really are. That is, the pyrogenic constituents represented 12 to 25 percent of the individual homologues that exceeded their individual MDL. Pyrogenic compounds would generally be lost in the noise in the presence of a strong petrogenic hydrocarbon signal.

4.5 Statistical Tests

4.5.1 Inferential Tests

Testing for significant differences in the sediment data set involved splitting the data into categories of interest and then running a randomization ANOVA or t-test to look for significant differences. For this report, we set the level of significance (alpha) to be 0.05 simply to be consistent with previous years' analyses (see recommendations).

Inferential testing of sediment data looked at two categories, i.e., spatial and temporal. Spatially, we were interested in seeing if there were significant differences among geographic categories. For these assessments, the sites were divided into three groups, Prince William Sound, Port Valdez, and Gulf of Alaska. The same sites were also tested in shallow and deep categories to evaluate the depth perspective. The spatial trends for sediments are presented in Table 4-13. Looking first at the depth groups, comparing the shallow versus the deep sites in cruises 6 or 8 when all sites were visited (n = 6,8), there were no significant differences in any index except TAHC in cruise 8 ($p = 0.04$). The finding of no-difference would not be an unexpected result considering the wide variance from the mix of reference and exposed sites within each depth group. But when the groups are split into their respective depths, all of the deep site tests and half of the shallow site tests had significant within-group differences. This suggests that sites are more similar within depth groups than between depth groups.

Within the Sound, the shallow Sheep Bay reference site (SHB-M) is significantly different in TPAH and CRUDE indices from the other three shallow sites in the PWS group, but results are mixed from the other indices. In the two deep PWS site comparisons, TPAH and UCM are significantly different, with mixed results from the other indices. In contrast, the two Port Valdez sites (Alyeska Marine Terminal and Gold Creek) are highly different for all indices except FFPI. Similar results are obtained comparing the tanker anchorage site, Knowles Head, to the Gold Creek reference site, GOC-S, but in this case only UCM is not different.

Outside of Prince William Sound, the Windy Bay site (WIB-M) appears to have fairly clean shallow sediments. It might serve as a shallow reference site for outside PWS except that when compared to Sheep Bay, five of the six indices were significantly different. Comparing Windy Bay to Shuyak Island (SHH-M) produced the same significant differences, only FFPI was not significant. Finally, at the deep sites, Aialik Bay (AIB-S) is unique in all but one index compared to both WIB-S and SHH-S.

Table 4-13. Randomization t-tests & ANOVAs testing spatial trends for sediment's oil indices

Sediments - Two Sample Randomization t-Tests

		Crude	TPAH	TAHC	UCM	CPI	FFPI
PWS		SHB-M vs DII-M		n=6,6 or 7,7			
		0.034	0.001	0.002	0.809	0.391	0.046
	Shallow	SLB-M vs SHB-M		n=6,6 or 7,7			
		0.013	0.011	0.111	0.003	0.576	0.002
		SHB-M vs KNH-M		n=5,5			
		0.009	0.005	0.009	0.438	0.733	0.415
		SHB-S vs DII-S		n=6,6			
		0.010	0.016	0.095	0.030	0.166	0.074
	Deep	SHB-S vs SLB-S		n=6,6			
		0.173	0.002	0.556	0.002	0.358	0.002
Valdez Arm		GOC-S vs AMT-S		n=9,9 or 10,10			
		0.000	0.000	0.000	0.000	0.000	0.546
		GOC-S vs KNH-A		n= 5,5			
		0.007	0.024	0.016	0.510	0.004	0.008

		Crude	TPAH	TAHC	UCM	CPI	FFPI
PWS/GOA		SHB-M vs WIB-M		n=5,5			
		0.009	0.008	0.053	0.008	0.031	0.834
GOA	Shallow	SHH-M vs WIB-M		n=5,5 or 6,6			
		0.009	0.002	0.002	0.003	0.012	0.169
		AIB-S vs SHH-S		n=5,5			
		0.007	0.010	0.010	0.586	0.024	0.009
	Deep	AIB-S vs WIB-S		n=6,6			
		0.003	0.003	0.002	0.033	0.091	0.004
		SHH-S vs WIB-S		n=5,5			
		0.416	0.009	0.009	0.080	0.005	0.014
Cruise 6		Shallow vs Deep		n=6,8			
		0.953	0.747	0.531	0.241	0.246	0.235
Cruise 8		Shallow vs Deep		n=6,8			
		0.141	0.415	0.043	0.873	0.515	0.626

One Way Randomization ANOVAs

Shallow Sites (less SLB)		
Crude 2	0.001	n=5x5 or 5x6
TPAH	0.000	
TAHC	0.000	
UCM	0.113	
CPI	0.052	
FFPI	0.085	

Deep Sites n=8x5	
Crude 2	0.003
TPAH	0.006
TAHC	0.000
UCM	0.005
CPI	0.000
FFPI	0.000

Control Sites AIB vs SHB vs GOC n=3x6	
Crude 2	0.000
TPAH	0.000
TAHC	0.003
UCM	0.118
CPI	0.000
FFPI	0.000

Lastly, an ANOVA was run to compare the three deep reference sites, AIB, SHB, and GOC. The results showed significant differences in all but the UCM index.
Temporally, we looked at annual sediment trends for each site (Table 4-14). Although the temporal statistics were slightly compromised by the small sample sizes (n ranged from 5 to 10 sampling events), overwhelmingly, there were no significant differences in any of the sediment oil indices at any site. The most significant results were reported from AMT-S for the CRUDE Index ($p = 0.08$) and the CPI ($p = 0.09$).

Focusing on the short term, we ran the same comparisons, but looked only at the sequence from cruise to cruise. For example, did TPAH go up at AMT between cruise 8 and 9? The t-test results are presented in Appendix VII. Overall, there were very few significant changes from cruise to cruise—23 hits in 456 tests using an alpha of 0.05.
Looking at just the last cruise for each site, several significant changes in the CRUDE index were observed. Sediment hydrocarbon loads increased at KNH-A and SHB-M and decreased at SLB-M. The Sleepy Bay shallow site also showed significant decreases in TPAH and UCM.

For the *Mytilus* tissue data, only 9 of 142 tests showed significant changes (alpha = 0.05) from the previous cruise. In the last sampling, cruise 9, the only significant changes identified were in TPAH at Windy Bay and MPI at Disk Island.
In summary, the inferential tests found no significant temporal trends across the cruises at any site (although CRUDE and CPI at AMT were only marginally insignificant, $p < 0.10$). The spatial tests suggest the sites are not changing in concert within their groups and in comparison to each other, except at the shallow sites within PWS where TPAH and CRUDE are changing similarly.

This battery of inferential statistics is only as good as the data allow. In a subjective assessment of time-series constituent histograms for a particular site, one may see changes in signature patterns reflecting a temporal trend. However, the trend cannot be statistically validated because either the embedded pattern cannot be isolated by the indices used for inferential statistics or the data series is too short to detect the trend. In either case, we have still reported the observed temporal changes as relevant to the objectives of this study.

4.5.2 Power Analyses

The power analysis compiled by KLI (1993) following the initial cruise, found that considering the small within-station variability, three-replicate sampling would be adequate to attain a target power of 0.80 at an alpha level of 0.05. In nonstatistical terms, that means that there is only a 20 percent chance of missing an effect combined with a 5 percent chance of incorrectly identifying an effect.

Table 4-14 Exact probabilities of representing a data set with random trends

STATION	Crude	TPAH	FFPI	TAHC	UCM	CPI
AIB-S	1.00	1.00	1.00	1.00	0.15	1.00
AMT-S	0.08	1.00	1.00	1.00	1.00	0.09
DII-M	1.00	1.00	1.00	0.51	0.51	0.17
DII-S	1.00	1.00	1.00	1.00	1.00	1.00
GOC-S	0.56	1.00	1.00	1.00	0.33	1.00
KNH-A	1.00	1.00	1.00	0.17	0.19	1.00
KNH-M	1.00	1.00	0.48	0.48	1.00	0.48
SHB-M	0.14	0.51	1.00	1.00	0.52	0.14
SHB-S	1.00	0.14	1.00	1.00	1.00	0.21
SHH-M	0.45	1.00	1.00	1.00	1.00	1.00
SHH-S	1.00	1.00	0.44	0.44	0.44	1.00
SLB-M	1.00	0.51	1.00	1.00	1.00	1.00
SLB-S	1.00	1.00	1.00	1.00	1.00	1.00
WIB-M	0.45	1.00	1.00	0.46	0.18	0.49
WIB-S	1.00	1.00	1.00	1.00	1.00	1.00

As a test of the initial three-replicate decision, we examined the results from every station and cruise in this study and calculated the power to detect a given difference. For sediments, we looked at mean TPAH, TAHC, UCM, FFPI, CPI, and CRUDE index values; for tissues, we assessed TPAH and FFPI only. Unfortunately, the randomization techniques we prefer to use do not have a statistical definition of power as do the parametric t tests and ANOVAs done by KLI. However, it is known that, in one special case, if the data set being analyzed perfectly fitted a normal distribution, the results and power of the parametric test would be exactly equivalent to those of the randomization test. As the data set deviates from normality, the parametric test results become more biased while randomization tests do not. This implies that the power of randomization tests is at least equal to that of parametric tests.

So, for this power validation exercise, we had to assume that parametric rather than randomization two-sample t tests would be run using the same parameters that KLI initially used (alpha = 0.05, n = 3) while looking for 25, 50, 75, or 100 percent differences. For example, the geometric mean and variance from the three replicate values for each oil index from a single site/cruise were used to calculate the power to detect 25-100 percent differences between the actual samples and a hypothetical additional set. The result of the calculations is the power of the statistic for each desired level of detection (a function of the data's variance or "spread" relative to the average value). The sediment results show that although there were a few lapses in detection power during earlier cruises, most of the cruise 9 samplings easily met the power target power of 0.8 (Appendix IV). From actual measured values, TPAH, TAHC, UCM (rather than the calculated indices), there were notable exceptions for TPAH at the shallow Windy Bay site (WIB-M-9, max power 0.433 at 100 percent difference), SHH-M-9 and KNH-M-9 at the 25 percent level. There were no problems with TAHC results from cruise 9, but UCM was highly variable at several sites (Appendix IV). The TPAH data from mussels met the target power in every cruise (Appendix V).

However, there are two caveats. These sterling results are based upon the assumptions of normally distributed data with equal variances. To achieve normally distributed data, the raw data must be either log or arcsine transformed, depending on the variable (KLI, 1993). This transform, of course, requires that the magnitude of detection also be computed from the transformed means.

The desired difference in means when back transformed to actual data values will thus be appreciably higher than expected. For example, the steep rise in tissue TPAH between AIB-B-8 and AIB-B-9 was 151.4 to 292.1 ng/g, an apparent 93 percent rise; however, after a log transform, the difference is actually only 16 percent. There is much less power to detect this magnitude of difference (0.86 at 16 percent vs. 1.00 at 25 percent) even though there is such a disparity in the raw data.

The other caveat lies with the assumption of equal variances. The same data run through a two-sample, two-tailed parametric t test assuming equal variances versus the same test with an assumption of unequal variances versus a randomization test assuming nothing about variances, produces 3 different results (Table 4-15). In this particular case (AIB-8 vs. AIB-9 TPAH), the variances are slightly unequal (F test, $p = 0.09$). According to the p values in the following table, none of the tests showed significant results for an apparent 93% spike in TPAH. The power assessments were based upon the equal variance method, which yielded the lowest probabilities and highest power.

Table 4-15. Probability results from various t-tests comparing methods of handling sample variance; AIB-B-8 vs. AIB-B-9 TPAH, n = 3,3.

Statistical Tests	One-tail	Two-tail
Randomization t test	0.083	0.202
Parametric t test with unequal variance	0.103	0.205
Parametric t test with equal variance	0.069	0.138

4.6 Program Comparability with Other Monitoring Programs

A good cross-validation of the RCAC program is to compare RCAC findings with other field programs. There were three comparable programs in Prince William Sound, the NOAA HAZMAT study, NOAA Mussel Watch program, and the compiled oil chemistry database from Trustee studies. Unfortunately, there is no overlap with the NOAA HAZMAT study. Their project composites sediment samples and only collects samples from a limited number of sites. Furthermore, they have drastically cut back on sampling in recent years.

Data are available from joint KLI/NOAA Mussel Watch sampling activities at Disk Island, Knowles Head, Sheep Bay, Shuyak Harbor, Sleepy Bay, and Windy Bay in 1995. These data are to be considered additional replicates rather than split samples. Additional Mussel Watch data are available from Mineral Creek Flats in the 1986 though 1995 time frame. Where appropriate, data from concurrent sampling periods with the LTEMP program for Mineral Flats were compared with profiles from Alyeska Marine Terminal and Gold Creek.

The Disk Island sample from the NOAA database was collected in March of 1995 at a site designated as PW-DI. The histogram plots and absolute concentrations of TPAH are very similar, suggesting good correlation between the two programs. Absolute concentrations are slightly different with an average value of 249 ng/g dry weight obtained for the LTEMP program and 605 obtained for the NOAA Status and Trends effort. In terms of composition, the histograms reflect input from weathered ANS oil along with background concentrations of pyrogenic PAH.

The Mussel Watch sample collected from Knowles Head in March 1995 also agrees very closely with the LTEMP sample. Absolute values of biphenyl, acenaphthalene, acenapthene, fluorene, pyrene, anthracene, dibenzothiophene, and numerous other PAHs are within the same range (1 to 10 ng/g dry weight), and the

only significant difference is higher concentrations of naphthalene in the LTEMP sample. Absolute concentrations for the LTEMP and NOAA samples are 99 and 130 ng/g dry weight, respectively.
NOAA data for Port Valdez were collected at Mineral Creek, reasonably close to the LTEMP Gold Creek station, in 1992, 1993, and 1995. In March of 1992 and 1993 and April of 1995, the NOAA Status and Trends samples show very light hydrocarbon loadings. The March 1992 was collected before the LTEMP program was initiated; however, very similar total aromatic loadings and histogram patterns were observed in March of 1993. Average concentrations for the NOAA Mineral Creek Flats sample were 782 ng/g dry weight and concentrations for the PWS RCAC station were 325 ng/g dry weight. Both samples show significant contributions from Alaskan North Slope crude oil. There are no NOAA data for the July 1993 sampling point; however, at this time, muscle data at the Alyeska Marine Terminal had dropped down to baseline levels. Although one would not expect an exact match for the two low-level samples taken at different sites and over a year time interval, there are similar TPAH levels at 163 and 248 ng/g dry weight for the NOAA and LTEMP samples, respectively. In terms of composition, the samples are slightly different with the Alyeska Marine Terminal sample showing higher levels of naphthalene, alkylated phenanthrenes, and extremely low-level (less than MDL) concentrations of dibenzothiophenes. Much better compositional agreement was obtained in the low-level samples at Mineral Creek Flats and Alyeska Marine Terminal when the background concentration levels were measured in 1992.
Therefore, while the match is not exact, the overall trends are at least very similar between the two programs. When concentrations increased again in 1994 from the *Eastern Lion* oil spill, NOAA Status and Trends sampling was not completed.

The comparability of the March 1995 LTEMP sample from Alyeska Marine Terminal and the April 1995 sample from Mineral Creek Flats is excellent. The same histogram pattern, showing evidence of Alaskan North Slope Crude oil is observed in both profiles, and the absolute amounts are very close with values of 672 and 517 ng/g dry weight reported for the NOAA and LTEMP samples, respectively. Once again, these data suggest excellent comparability between the two monitoring efforts.
The results for inter-program comparability at Sheep Bay are also extremely good. The histogram profiles are essentially matching while the absolute hydrocarbon concentrations are extremely close, 59 and 66 ng/g dry weight for the NOAA and RCAC LTEMP programs, respectively.
At Shuyak Harbor, the March 1995 samples are identical between the two programs. Histogram plots are consistent with exactly the same set of analytes being reported, and the TPAH values were 66 and 59 ng/g dry weight for the NOAA and LTEMP programs, respectively.
In comparing the samples between the two programs collected at Sleepy Bay, there are some subtle differences in the histogram profiles. Specifically, the NOAA sample contains higher levels of C2- and C3-fluorenes; however, the other aromatics are very consistent. TPAH concentrations of 857 and 623 ng/g dry weight were obtained for the NOAA and LTEMP programs, respectively.

The last samples for which replicate data were taken occurred at Windy Bay. In this case, the histograms are very similar with the same analytes being detected in both samples. These analytes are indicative of soot from burned oil or the ubiquitous and extremely low-level background signal observed elsewhere throughout Prince William Sound. Absolute concentrations between the two programs were very comparable with TPAH values of 53 and 62 ng/g dry weight obtained for the NOAA and LTEMP programs, respectively. Table 4-16 compares the results from the two programs when complementary samples were collected.

Table 4-16. Comparison of Data from LTEMP and NOAA Mussel Watch Tissue Analyses.

Mytilus Collection Site	LTEMP (ng/g)	NOAA MW (ng/g)	Noted Differences
Disk Island	249	605	
Knowles Head	99	130	higher concentrations of naphthalene in the LTEMP sample
Sheep Bay	59	66	
Sleepy Bay	857	623	NOAA sample contains higher levels of C2- and C3-fluorenes
Windy Bay	53	62	

Another available data set is the Trustee's oil chemistry database. These data are heavy with samplings in the early post-spill period prior to the start of RCAC's study. However, there are three general locations that match with RCAC sites, namely, Disk Island, Sleepy Bay and Windy Bay. At these sites both shallow and deep sediments were collected and analyzed in a comparable manner. Unfortunately, the sampling dates precede RCAC's study at Disk Island and Windy Bay, but the data are relevant to demonstrate a reduction in hydrocarbons since 1989.

The data from Sleepy Bay are highly pertinent. The Trustee's study sampled 3 replicates at a reported 40 m station on 11 July 1993; KLI sampled at 35 m on 22 July 1993. The CRUDE index values are very similar (Trustee's 110 +/- 27, RCAC 144 +/- 2.5; $p = 0.104$) and certainly fit the trend of reducing hydrocarbons at the site. Figure 4-13 presents the Crude index values calculated for sediments collected at Disk Island, Sleepy Bay, and Windy Bay during the two programs. As noted above, there is only one station where exact overlap in sampling occurred (Sleepy Bay July 1993), but in that case, the plotted CRUDE index values are nearly identical. In other locations the CRUDE index values calculated with data from each program complement each other by demonstrating similar trends or comparable values over time.

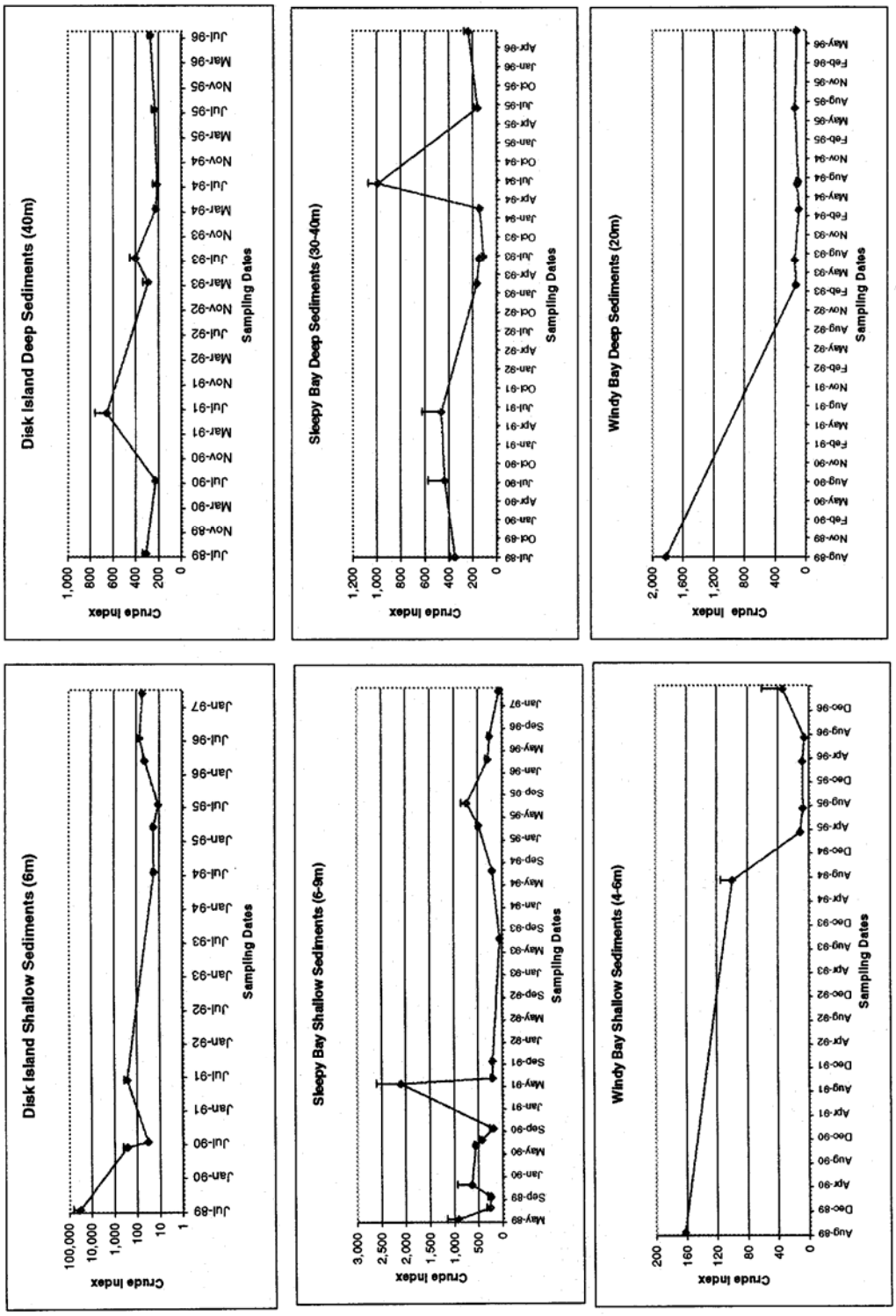


Figure 4-13. Comparison of TPAH values from Trustee studies at Disk Island, Sleepy Bay and Windy Bay locations prior to RCAC samplings (begun Mar 93) with later RCAC values. A single sampling overlap occurs between studies at Disk Island, deep sediments on July 21-22, 1994. Error bars represent SE of mean.

4.7 Analytical Chemistry Issues

4.7.1 Method Detection Limits

The number of hours to investigate this issue is beyond the scope of this report, but we were concerned by the fact that MDLs (particularly for the PAHs) changed so dramatically between cruises. Also, the MDLs seemed, on average to go down from cruises 1 and 2 to significantly lower levels for cruises 3, 4, 5, and 6, and then increase again for cruises 7 and 8, only to finally drop to even lower levels for cruise 9.
Rather than address these apparent changes on a compound specific basis, we chose to mark the average MDL value on the side of every histogram as a visual reference. All of the histograms presented and discussed in Appendix I have been so-marked to aid in assessing overall patterns and easily recognize when individual analytes were above or below the average MDL.

4.7.2 Procedural Blanks

An analytical procedural blank is run with each batch of 20 sediment and mussel samples analyzed during the program. These blanks from laboratory stock consist of HPLC-grade water that is run through each step of the analytical procedure. The blanks essentially represent a combined lab ware, solvent, and instrument check to evaluate artifacts that might be generated by the "analytical system." They also help to assess "carry over" of higher-level sample contamination from one sample to another.
Usually, procedural blanks generate signals that are well below the target analyte signals observed in field samples; however, in this program, the field PAH levels are so low (especially for mussels and shallow sediments) that special concern is required in evaluating laboratory contaminants. The PWS RCAC protocol states that, "If blank levels for any component were greater than three times the MDL, the procedure and instruments were to be investigated to identify sources of contamination." Furthermore, any batch of samples whose associated blank failed these criteria must be rerun. It is not clear from the data sets examined just how often it was necessary to rerun sample batches; however, as shown below, we do have some concerns about data interpretations for mussel and sediment samples obtained from extremely clean areas.
Figure 4-14 presents a typical summertime compositional pattern for low-level PAH concentrations in a mussel tissue sample. In this case, there is evidence of fairly high naphthalene concentrations; however, there are no phenanthrenes, dibenzothiophenes, or chrysenes. The total PAH concentrations in the sample are 77 ng/g dry weight. As in most of the mussel samples collected at this site (and elsewhere in Prince William Sound), there is evidence of higher molecular-weight pyrogenic compounds between benzo(b)fluoranthene and benzo(g,h,i)perylene.

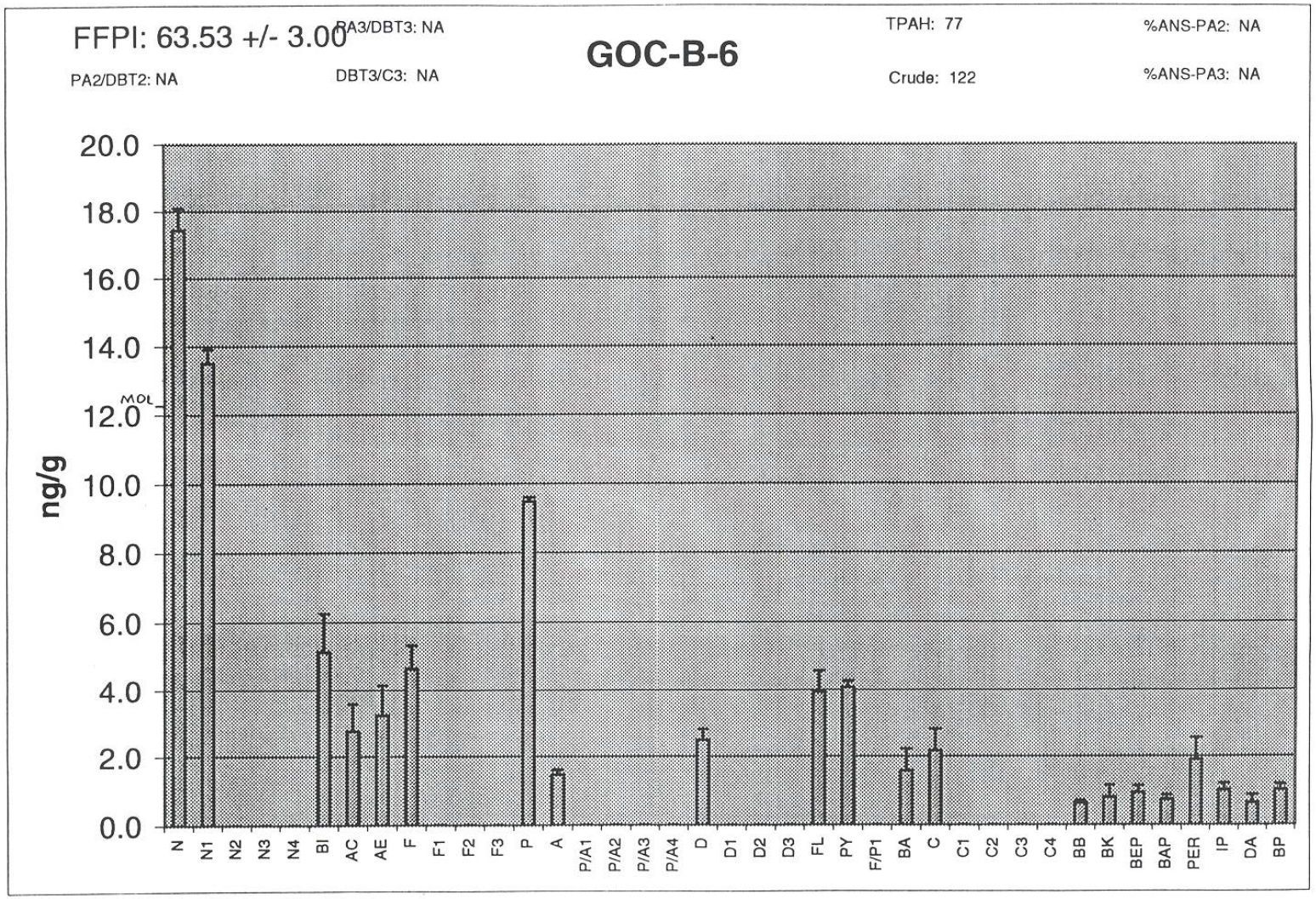


Figure 4-14. PAH Histogram for Gold Creek Mussel Tissue Sample, Cruise 6 (July 1995)

This pattern, or one very similar to it, presents a bit of a dilemma. It was observed very often in the cleaner areas where few, if any, alkylated PAH derived from the more common oil sources were observed. Many of the constituents (naphthalene, fluorene, phenanthrene, anthracene, fluoranthene, pyrene and the benzo(b)fluoranthene through benzo(g,h,i)perylene suite) have been identified by NOAA (1997) as combustion derived PAH by-products from burning oil. At the same time, however, Bence and Burns (1995) describe a similar profile as "procedural artifacts." In actual fact, many of these components in this exact pattern do routinely show up in procedural blanks run at GERG. The levels are quite low, usually well below the statistically defined method detection limit, and usually an order of magnitude lower than their respective concentrations in most field samples. Nevertheless, as shown by the procedural blank shown in Figure
4-15, it is extremely disconcerting to see this same pattern repeat so often in low-level samples.
In this example the TPAH level in the blank is only 3.4 ng/g; however, blank levels ranging from 1.9 to an extremely bad blank at 97 ng/g (assumed dry weight) were reported in the QC data set generated with this program (Table 4-17). Figures Proc Blk Q13739, Q13789, Q13739 and Q13783 in Appendix VI are illustrative of the types of Procedural Blanks that have been reported.
Because of the ubiquitous nature of this pattern and components which are common to both combustion products (as identified by NOAA) and procedural artifacts (as defined by Bence and Burns), we have elected to call this pattern "Combustion by-products/artifacts" in this report.

4.8 Field Sampling Issues

4.8.1 Field Blanks and Equipment Rinsate Blanks

Field blanks were collected once during the collection of each 20 replicate samples collected by grab on each survey for the analysis of PAH and AHC. Field blanks consisted of HPLC-grade deionized (DI) water poured from the DI water container into the appropriate sample container. Field blank analyses were used to assess the accuracy and comparability of the data by determining if contaminants such as engine exhaust were present during sampling.
Equipment reinstate blanks were also collected once per 20 grab samples for the analysis of PAH and AHC. Equipment blanks consisted of a deionized water rinse of the grab after it had been decontaminated. By providing information to determine if the decontamination process was adequately cleaning the grab and sampling utensils, these blanks helped to assess whether the data were comparable and representative. A procedural blank consisting of the HPLC water was run with each batch of field blanks for PAH and AHC.

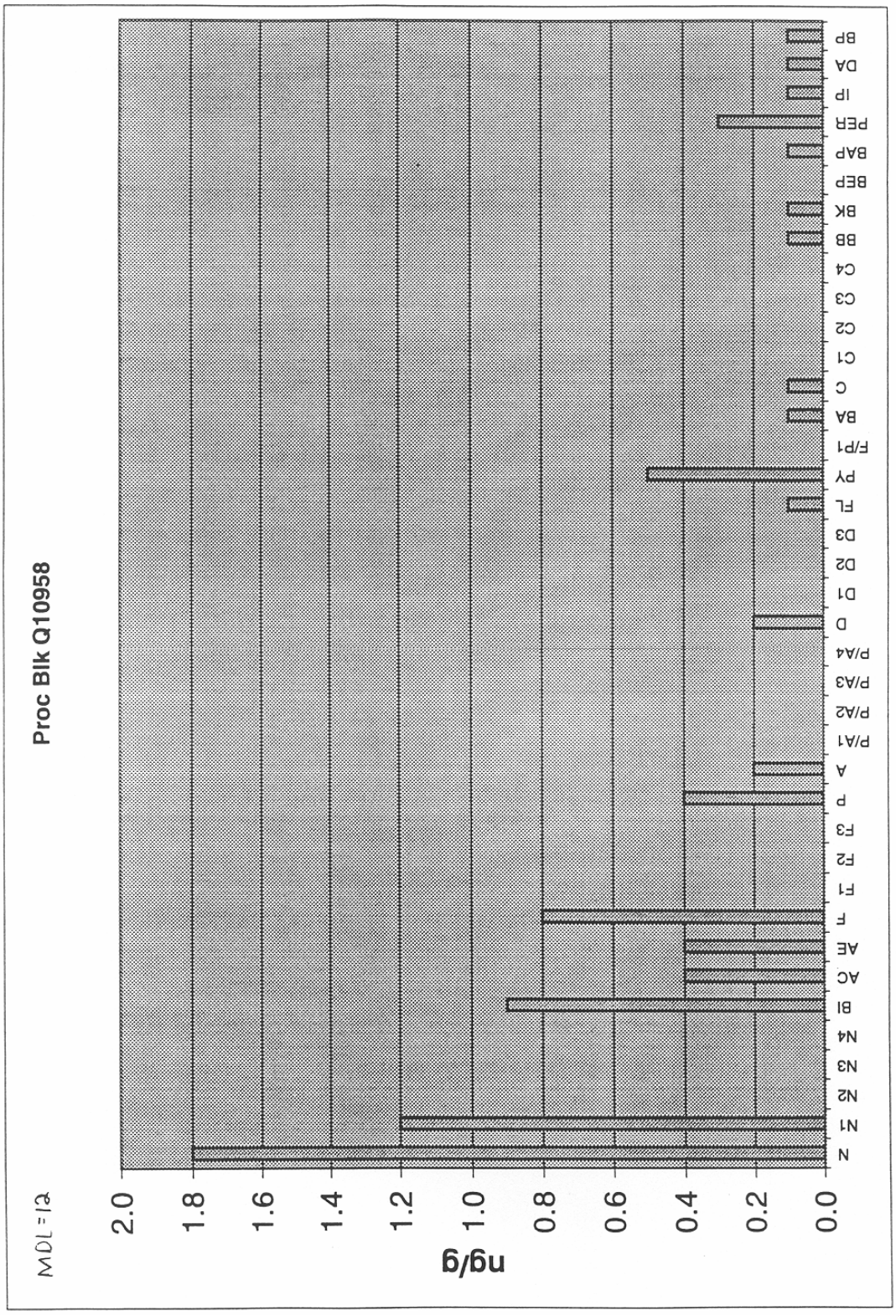


Figure 4-15. PAH Histogram Plot from Laboratory Procedural Blank, August 1995.

Table 4-17. Concentrations of Hydrocarbon Constituents in Laboratory Procedural Blanks

LAB ID	TPAH	TAHC	UCM	Extr Date
Q5112	1.79	1	0	4/9/93
Q5388	1.9	14	0	5/18/93
Q5415	5	71	0	5/19/93
Q5421	3	42	0	5/19/93
Q5436	4.6	16	123.3	5/21/93
Q5444	3	12	0	5/22/93
Q5454	4.4	71	0	5/26/93
Q5473	4.3	41	0	5/28/93
Q5506	3.5	53	0	5/31/93
Q5951	3.5	4	0	8/5/93
Q6034	6.8	8	0	8/19/93
Q6287	5.2	32	0	8/20/93
Q6039	10.3	7	0	8/20/93
Q6284	4.5	32	0	8/21/93
Q6042	2.9	6	0	8/23/93
Q6861	4.3	29		10/13/93
SDS1_Q639	7	18	0	10/15/93
Q7144		58	0	11/29/93
SDS2_Q6393		64	0	1/17/94
Q9065	27.6	73	3.3	5/17/94
Q9215	3.9	19	9.2	6/8/94
Q9225	7.7	113	0.3	6/9/94
Q9239	3.4	29	0	6/13/94
Q9229	12.4	254	103.7	6/13/94
Q8534		231	0	7/8/94
Q9659	31.65	726	0	8/3/94
Q9670	2.7	32	0.7	8/5/94
Q9675	2.9	51	0	8/8/94
Q8690	12.6	263	0	8/9/94
Q9702	1.9	408	0	8/11/94
Q9711	4	48	0.2	8/16/94
Q9725	13.9	477	0	8/17/94
Q9810	2.8	374	0	9/8/94
Q10438	8.47	23	0	4/24/95
Q10458	7.8	393	0	4/29/95
Q10505	5.3	0	0	5/10/95
Q12101	3.4	17	0	5/24/95
Q12107	3.6	5	0	5/26/95
Q12635	3.5	31	2.2	8/2/95
Q12646	8.2	21	0	8/10/95
Q12658	11.1	11	0.8	8/14/95
Q10958	7.5	0	0	8/16/95
Q10964	12	0	0	'8/17/95
Q12697	9.9	38	0	8/18/95
Q12714	4.5	37	0	8/22/95
Q13739	58.8	233	0	4/9/96
Q13783	6.2	57	0.7	4/17/96
Q13789	3.5	47	0	4/19/96
Q11563	14.4	0	0	4/24/96
Q11569	13.4	0	0	4/29/96
Q11655	9.5	0	0	5/14/96
Q11864	50.3	266	0	8/7/96
Q14241	6.8	0	0	8/22/96
Q11886	6.9	0	0	8/30/96
Q14255	3.3	32	0.3	8/30/96
Q14261	6.8	55	2.9	9/3/96
Q11899	9.8	0	0	9/5/96
Q14267	4.5	23	0	9/5/96
Q14273	5.7	33	0	9/9/96
Q12010	5.3	0	0	10/4/96
Q12301	97	0	0	2/18/97
Q15044	22	257	0	3/21/97
Q15025	1.7	9.4	0.4	4/1/97
Q12381	19.8	0	0	4/2/97
Q15035	6.2	7.9	0.5	4/9/97
Q12451	85.7	0	0	5/13/97

Numerous problems were uncovered during the analysis of the field blanks completed in this program. For example, Figure 4-16 shows a field blank collected from Sheep Bay during July 1994. In this case, the aliphatic fraction shows a small suite of diesel-like n-alkanes ranging from n-C10 through n-C 24 with a maximum at n-C15. The aromatic fraction shows the same suite identified by NOAA as being derived from oil combustion, including the complete suite of higher-molecular-weight aromatics in the benzo(b)fluoranthene through benzo(g,h,i)perylene range. Their relatively greater distribution compared to the lower-molecular-weight components may suggest a source associated with lube oils or greases used on the research vessel.

Seven other examples of contaminated field blanks are presented in Appendix VI ranging from sources such as combustion products (probable diesel soot), lube oil or grease, and cross contamination from dirty sediment samples themselves.

In examining all of the available field and equipment rinsate blanks generated during the entire program, not one sample was observed that did not represent some type of contamination. Table 4-18 lists the TPAH and TAHC values obtained for the all the field and equipment rinsate blanks collected during the program. As shown by the data in that table, TPAH concentrations ranging from 8 to 120 ng/g were observed and TAHC levels generally fell between 1.4 and 373 ng/g. While these concentrations are generally below the TPAH and TAHC levels observed in the actual field samples, it is clear that there are concerns with potential sample contamination in the field. We acknowledge that a contaminated equipment rinsate blank may or may not indicate a serious problem (i.e., the sediment samples are taken from the middle of the grab sample where the sediment isn't in contact with the walls of the sampler). As such, the rinsate blank may not represent the true exposure of the sample, but the potential for contamination still exists.

Considering the above information, the RCAC project mangers have a decision to make. Are they concerned enough about the low-level PAH/AHCs to require from the contractors, tighter controls over sample and procedural contamination? Or is pursuit of the low-level hydrocarbons beyond the objectives of the program and not worth the extra effort? Is it enough to just know that hydrocarbons may exist but are below the limits of MDL and blank contamination?

In either case, the data suggest that the contractors should be advised that blank contamination, and presumably sample contamination, is at a threshold (albeit low) that affects the qualitative interpretations and statistical analyses of the field data. Greater awareness should be directed to ensure that contamination from combustion sources or physical contact of sampling equipment with oily products is avoided.

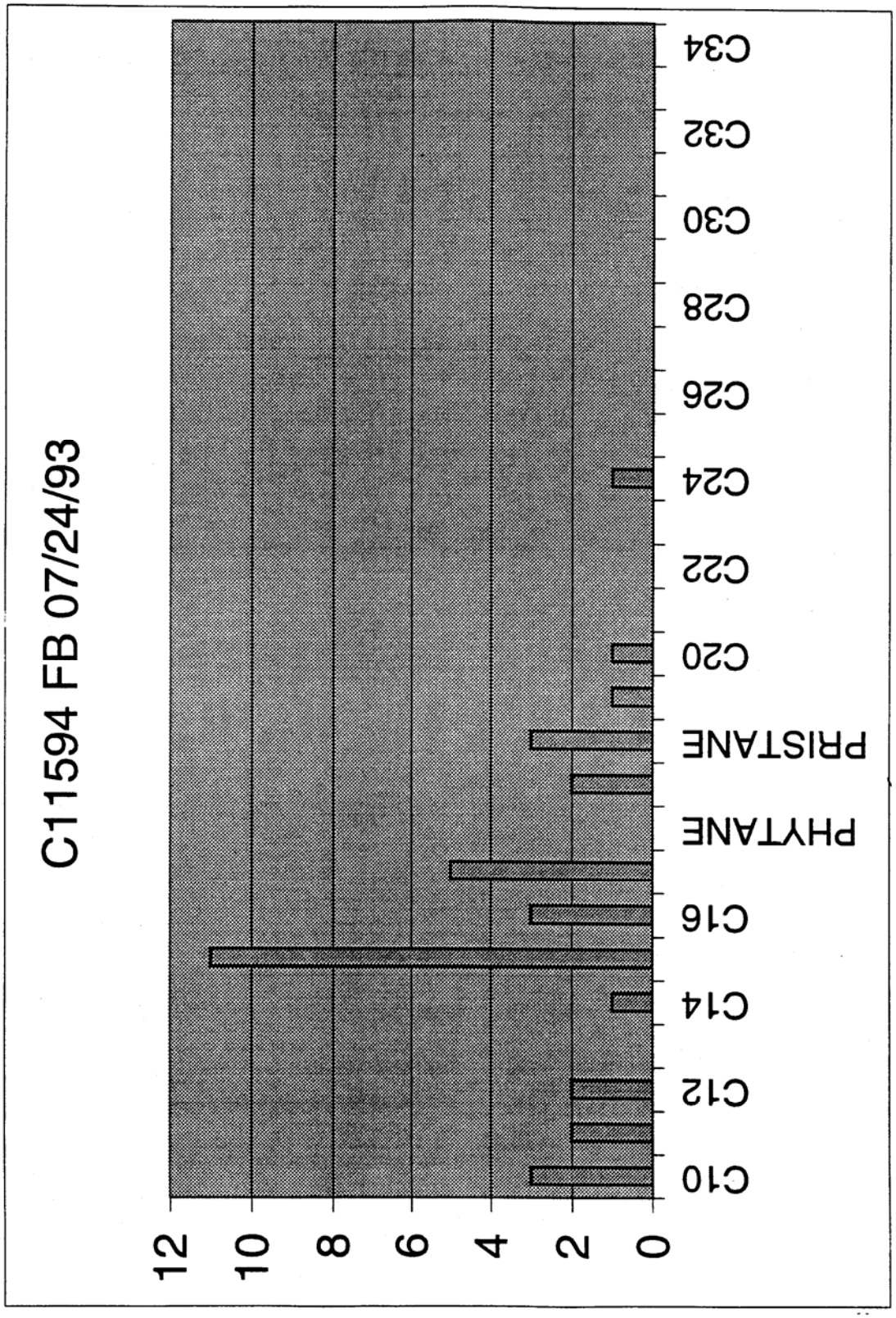


Figure 4-16a. Histogram Plot from Aliphatic Fraction of Field Blank Generated at Sheep Bay, July 1993.

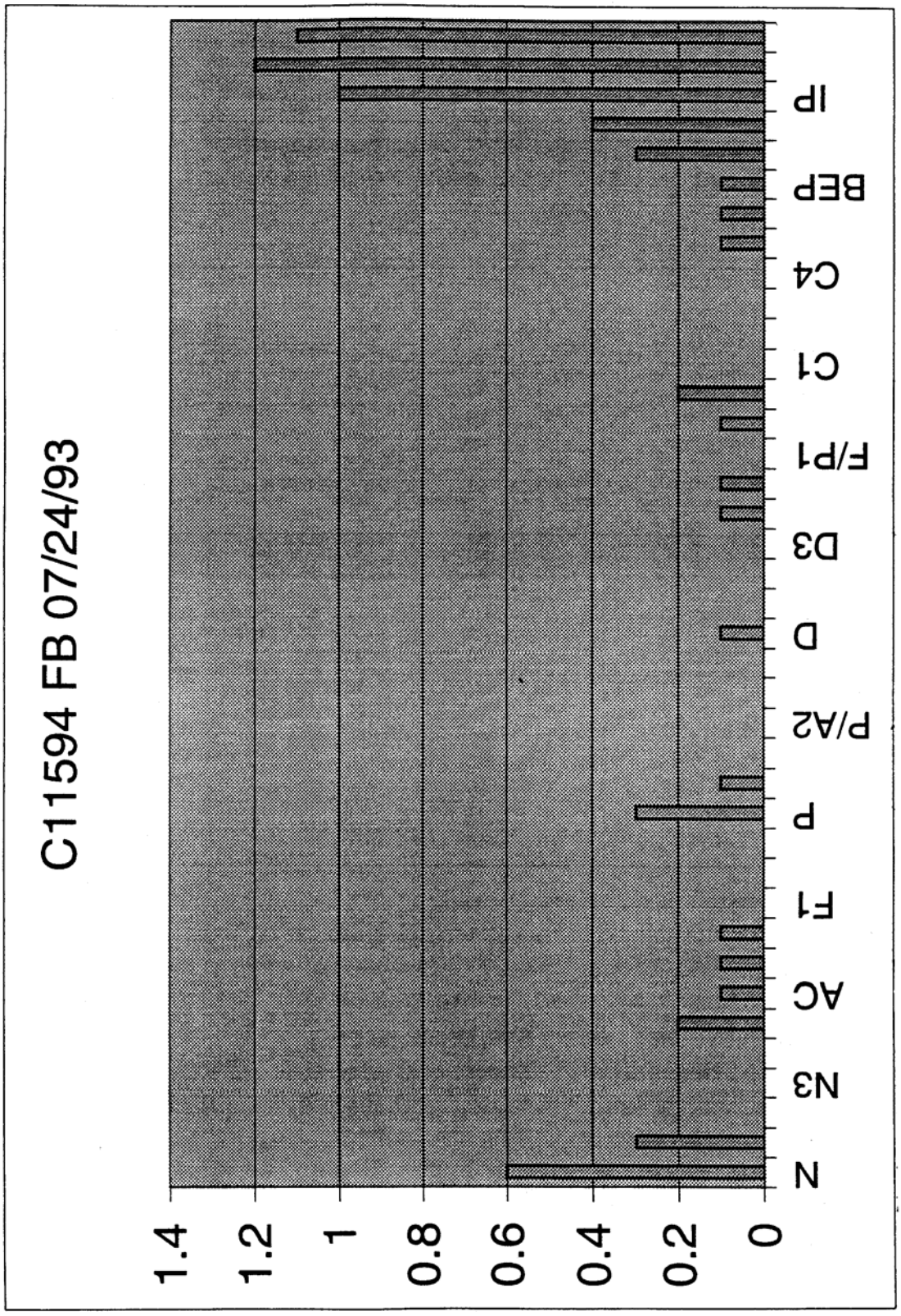


Figure 4-16b. Histogram Plot from Aromatic Fraction of Field Blank Generated at Sheep Bay, July 1993.

Table 4-18. TPAH and TAHC from equipment and field blanks.

SampType	LabSampl	StnID	SampDate	TAHC	TPAH
Equip Blar	C10851	KNH-S-1	3/20/93	26	1.43
	C10855	SLB-S-1	3/24/93	27	1.77
	C10859	AMT-S-1	4/3/93	55	2.52
	C11593	WIB-S-2	7/20/93	26	8.2
	C11595	SHB-S-2	7/24/93	120	4.4
	C13945	SHH-S-3	3/22/94	59	373.2
	C13947	GOC-S-3	3/26/94	57	66.8
	C14737	DII-S-4	7/22/94	26	5
	C14837	SHH-S-4	8/1/94	57	2.9
	C19764	GOC-S-5	4/3/95	40	11.71
	C21482	WIB-S-6	7/13/95	16	8.5
	C21485	KNH-A-6	7/18/95	18	30.7
	C23614	AMT-S-7	3/16/96	29	32.1
	C25321	SLB-S-8	7/14/96	25	4.3
	C25323	KNH-A-8	7/21/96	20	5.3
	C26471	GOC-S-9	3/6/97	44	33.9
Field Blanl	C10853	KNH-S-1	3/20/93	13	1.91
	C10857	SLB-S-1	3/24/93	8	1.99
	C10861	AMT-S-1	4/3/93	11	2.3
	C11592	WIB-S-2	7/20/93	27	5.8
	C11594	SHB-S-2	7/24/93	35	6.3
	C13944	SHH-S-3	3/22/94	19	44.9
	C13946	GOC-S-3	3/26/94	16	43.6
	C14739	DII-S-4	7/22/94	42	2.5
	C14836	SHH-S-4	8/1/94	19	2.2
	C19766	GOC-S-5	4/3/95	20	9.97
	C21483	WIB-S-6	7/13/95	18	3.2
	C21484	KNH-A-6	7/18/95	32	3.5
	C23616	AMT-S-7	3/16/96	10	13.8
	C25320	SLB-S-8	7/14/96	27	4.8
	C25322	KNH-A-8	7/21/96	27	4.5
	C26470	GOC-S-9	3/6/97	31	30.9

4.8.2 Sediment Depth Issues

The range of on-site sampling depth appears to be correlated with the reported data values at some sites in the program. Of obvious concern is the steep bottom topography in Port Valdez fjord where it is certainly a difficult place to sample and where the high depth variance might be expected to cause problems in the data set (i.e., a change in depth suggests a change in depositional regimes). However, a more severe data correlation with depth occurs at the shallow sampling sites where just a couple of meters difference in depth can move the sampler into a different depositional regime. The data are insufficient to fully assess the issue; however, the hydrocarbon (TPAH, TAHC, CRUDE index) and grain size data from shallow samples at Disk Island and both shallow and deep samples at Sheep Bay appear to be strongly correlated with depth. This depth dependence is illustrated in Figure 4-17. These issues are addressed further in Appendices 1 C and 1 F.

4.9 Secondary Objective - Identify, Define, and Describe Environmental Variables Potentially Affecting Hydrocarbon Concentrations.

A variety of factors and mechanisms have the potential to influence hydrocarbon concentrations in water, sediments, and biota. Generally, the types of operative factors are different in sediments and animals. Hydrocarbon concentrations are more strongly influenced by physical and microbial factors in sediments and biochemical and physiological factors in animals.

4.9.1 Intertidal and Subtidal Sediments

In intertidal and shallow subtidal sediments, important factors influencing hydrocarbon concentrations include: weathering rates, as mediated by insulation, air and/or water temperature; the nature and activity level of the microbial flora; inorganic nutrient concentrations in sediment and water; transport and dilution rates, as mediated by wave action, tidal flushing, currents, emersion/immersion cycles, circulation (exchange) rates; sediment type and texture; and sediment mixing and bioturbation rates.

In intertidal and shallow areas, wave action is a more important factor than in deeper water. Moreover, higher water temperature, light, DO, and inorganic nutrient levels probably result in more rapid rates of degradation. In addition, transport mechanisms and gravity will typically carry hydrocarbons out into deeper water. Deep sediments are often sinks into which contaminants become deposited. Weathering rates may be slow because of reduced temperature, nutrient availability, and light.

According to Roberts et al. (1996), weathering rates in sediments are inversely related to the concentration of the oil. They stated: “The weathering of oil and concentrations detected are directly related. Those sites in PWS with reduced degradation still contain the highest petroleum concentrations of “persistent oil” when it is environmental factors that created this persistence. The current monitoring sites indicate that the most persistent surface samples were found in highly sheltered locations, exposed to less physical

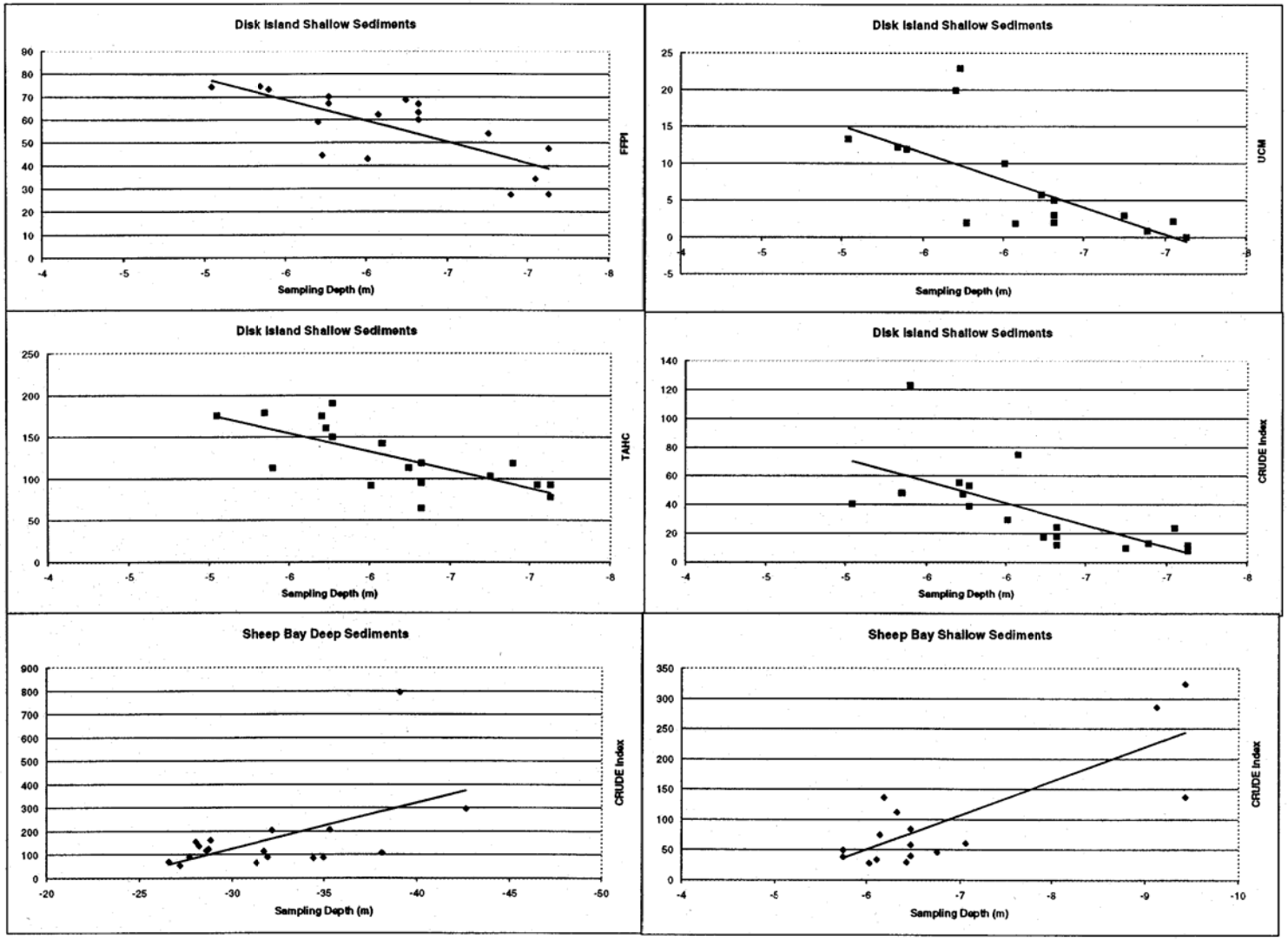


Figure 4-17. Comparison of sampling depth variance versus variance of selected sediment variables from Disk Island and Sheep Bay, Cruises 1-9. Any deviation from a flat trend suggests the sediment variable may be confounded with sampling depth.

Reworking, but degradation was still possible by photo-oxidation and microbial activity. The subsurface samples, exposed to less reworking than surface sediments and less photo-oxidation, contained the most significant concentrations of oil. The principal weathering process for these subsurface samples is microbial degradation with possible sediment reworking during storm events." Griffiths et al, (1984) observed that, except at very low levels, rates of microbial degradation in oiled sediments were inversely related to the concentration of hydrocarbons. If true, low concentrations of oil (100 to 500 µg/g) are microbiologically degraded more rapidly than higher concentrations. However, samples at trace concentration (less than 1 µg/g) do not appear to follow this trend, probably due to environmental influences.

4.9.2 *Mytilus* Tissues

A valid concern when examining the results of a mussel watch program is that levels of petroleum hydrocarbons in mussels might not represent environmental levels due to physical, biological, or physiological phenomena. Physical factors that influence hydrocarbon concentrations in mussels include solubility of the oil, interactions between hydrodynamic energy levels and sediment type, and the rate at which the water-soluble fraction, suspended particulates of oil, and contaminated sediment are produced and become available for assimilation by the mussels. Important biological and physiological factors include: rates of feeding, bioaccumulation, depuration, metabolism and detoxification, and preferential uptake (i.e., the ability of filtering or deposit-feeding organisms to select or avoid substances or particles of a specific size). Houghton et al. (1992) reported on rates of bioaccumulation and depuration in resident and transplanted mussels at Smith Island and Eshamy Bay. They found that bioaccumulation and depuration occurred rapidly. These findings agree with those of NOAA (1989) and Pruell et al. (1987).

5. Conclusions and Integration of Detailed Site Observations

In the 4.5 years of data collection for this program, there have been some changes in the sampling design regarding station coverage, particularly for deep and shallow sediments. As a result, sediment data are not always available for all stations at both depths and both seasons. In contrast, mussel tissue collections have been essentially complete for the March and July samplings at all but one of the stations. Sample coverage is sufficient to allow identification of temporal trends and sources of individual hydrocarbons at the different stations.

Hydrocarbon levels in both sediments and tissues are generally very low. Average TPAH concentrations in deep sediments range from less than 30 ng/g dry weight at the cleaner sites (e.g., Aialik Bay) to around 300 ng/g dry weight adjacent to the Alyeska Marine Terminal in Port Valdez, to over 500 ng/g dry weight at Sleepy Bay. Shallow sediments generally have much lower TPAH levels, ranging from less than 10 ng/g dry weight at Knowles Head to approximately 400 ng/g dry weight at Sleepy Bay. Average TPAH hydrocarbon burdens in the tissues ranged from 130 ng/g dry weight at Aialik Bay to around 510 ng/g dry weight at Alyeska Marine Terminal. Compared to results from mussels analyzed in the vicinity of active ports and at oiled sites following the *Exxon Valdez* oil spill, these levels are generally quite low (Houghton et al. 1992; Roberts et al. 1996).

TPAH patterns in the sediments and tissues examined in this program reflect several "background" sources, including the Katalla oil seeps and coal particles derived from the rivers east of Prince William Sound, as well as oil-transportation activities associated with the Alyeska Marine Terminal in Port Valdez. At this time, there is a debate in the scientific literature as to whether the natural "background" hydrocarbons are actually derived from oil seeps or from coal particles. However, for the purposes of this program, it is sufficient that these fingerprints can be distinguished from the pattern generated by Alaskan North Slope crude oil introduced from present-day activities or from weathered EVOS residues.

Table 5-1 presents an integration of the more salient observations extracted from the entire data set collected over the 4.5 years of the program. This table was generated after detailed examination of every aliphatic and aromatic hydrocarbon histogram plot for every sample, consideration of the graphical trend analyses data from the CRUDE and *Mytilus* Petrogenic Indices, and calculation of relative ratio data for selected components that are characteristic of different oil sources. From the data presented in the table, it

is possible to identify stations in which high variability or hot spots indicating higher oil concentrations were noted over time. A subset of those hot spots was then further examined for evidence of ANS or EVOS oil, and these stations and cruises are also identified in the table.

Table 5-1. Summary of Sediment and Mussel Tissue Hits and Percent ANS / EVOS Crude Contributions for Highly Variable and High Oil Concentration Samples -- LTEMP Cruises 1-9, March 93 - March 97.

Site	Matrix	Temporal Changes in CRUDE Index	Overall Range of Average CRUDE Index Values		TPAH Concentration (ng/g DW)		TAHC Concentration (ng/g DW)		Cruise No. for Hot Spots or High Variability in CRUDE Index	Cruise No. for Positive Hits on ANS or EVOS oil[1]	Range of ANS or EVOS Contribution to TPAH Burden[2] (% of TPAH)		Influence of Depth Within Sampling Regime[3]	Crude Index Value from Last Cruise[4]
			Min	Max	Mean	C.V.(%)	Mean	C.V.(%)						
AIB	Deep	No	15	38	26	36	235	17		None	n.a.		CRUDE	38
	Shallow													
	Mytilus	Possibly	35	231	127	61			3,8,9	None				231
AMT	Deep	Yes	365	2,230	330	66	1810	22	All (1-9)	All (1-9)	25	54		762
	Shallow													
	Mytilus	Yes	56	1,500	512	89			3,4,5,9	1,3,4,5,7,8,9	49	100		496
DII	Deep	No	210	402	200	33	615	21						272
	Shallow	No	11	71	18	31	125	25		None			CRUDE	56
	Mytilus	Yes	55	742	250	92			3,4,5,9	3,4,5,8,9	30	82		290
	Intertidal				105,000	6	6360	110	6,8	6,8	93	100		
GOC	Deep	No	22	49	53	31	610	29		4	95			38
	Shallow													
	Mytilus	Yes	49	530	386	70			1,3,4,5,9	1,3,5,9	49	71		331
KNH	Anc. Sed.	No	63	201	136	59	310	57		None				201
	Shallow	No	10	25	8	16	45	31		None			TAHC	11
	Mytilus	Yes	52	412	243	65			3,4,8,9	8	50			412
SHB	Deep	Yes	79	402	119	41	432	47	8,9	None			TPAH	402
	Shallow	Yes	30	249	64	46	309	66	8,9	None			CRUDE	249
	Mytilus	Yes	48	318	178	72			2,4,8,9	None				318
SHH	Deep	No	142	363	240	10	693	31		3	99 (n=1)			143
	Shallow	Possibly	56	135	76	41	326	19	8,9	None				113
	Mytilus	Possibly	34	289	136	89			8,9					268
SLB	Deep	Yes	142	993	510	112	373	23	4	None				237
	Shallow	Yes	68	724	412	49	823	64	5,6,7	4,6,7	11	100	CRUDE	67
	Mytilus	Yes	52	1,690	487	137			3	1,3,4,5,9	23	54		245
WIB	Deep	No	92	144	158	23	2160	11		3,4,6,8	21	49		123
	Shallow	Possibly	6	34	15	91	158	30	9	None			TAHC	34
	Mytilus	Yes	33	463	144	110			9	9	20			462

Notes:
1) Short and Babcock (1996) multiple ratio approach (minimum total chrysene criteria set at 10 ng/g dry weight)
2) Page et al. (1995, 1996) C2-Phenanthrene/C2-Dibenzothiophene ratio approach.
3) Entry denotes possible minor confounding influence of depth on noted parameter within the respective shallow or deep sediment regime.
4) Value for last sample collected (Cruise 9 for Mytilus, 8 for deep sediments, and 9 for shallow sediments).

From these analyses, the following distribution of ANS or EVOS-related oil was observed:

- Alaska Marine Terminal – ANS oil was detected in deep sediments during all nine cruises; mussels showed evidence of ANS oil for cruises 1, 3, 4, 5, 7, 8, and 9.
- Disk Island – No significant levels of ANS or EVOS oil were detected in any deep or shallow sediment samples; mussels exhibited evidence of EVOS oil in cruises 3, 4, 5, 8, and 9; intertidal sediments (opportunistically collected when weathered oil was observed on the beach during mussel collection) showed significant quantities of EVOS oil in cruises 6 and 8.
- Gold Creek – ANS oil was observed only once in deep sediments during cruise 4; mussels showed ANS oil in cruises 1, 3, 5, and 9.
- Knowles Head -- no evidence of ANS crude was noted in either the anchorage or shallow sediment locations; however, ANS crude was detected in mussel tissue samples collected during cruise 8.
- Sheep Bay -- no samples showed any evidence of ANS or EVOS oil for either sediments or mussel tissue.
- Shuyak Harbor -- ANS or EVOS oil was noted in the deep sediment for cruise 3 only; however, it was only detected in one of three replicates; mussel tissues showed higher concentrations of petroleum hydrocarbons in cruises 8 and 9, but they could not be positively identified as ANS or EVOS oil.
- Sleepy Bay -- no evidence of ANS or EVOS oil was observed in the deep sediment samples; however, shallow sediments showed positive hits for ANS or EVOS-derived oil during cruises 4, 6, and 7. Mussel samples showed evidence of ANS or EVOS oil in cruises 1, 3, 4, 5, and 9.
- Windy Bay -- ANS or EVOS-derived oil was observed in the deep sediments during cruises 3, 4, 6, and 8, but not in any of the shallow sediments. ANS or EVOS oil was detected in the mussel samples only during cruise 9.

Table 5-1 also lists the range of values obtained for the relative ANS or EVOS contribution to the total aromatic hydrocarbon burdens in the different sample matrices at the different stations. The relative contributions range from nondetect to 100 percent of the total PAH measured in a few samples. It should be remembered, however, that TPAH levels in most of these samples were extremely low. Therefore, although the percent ANS or EVOS oil for any given station may have been high, the absolute value for the concentration of residual oil itself was extremely low.

On initial examination, the hydrocarbon patterns observed for the mussel tissues may appear exceptionally variable with no apparent trend or explanation. The trends observed in Figure 4-4 (the *Mytilus* Petrogenic Index plot, page 38) can be correlated, however, with a chronology of documented events that have occurred within Port Valdez and Prince William Sound since 1993.

In May 1994, the *Eastern Lion* oil spill occurred at the Alyeska Marine Terminal during loading operations. Mussel samples collected at the time of the spill from the Alyeska Marine Terminal station (located near berth 5, the site of the spill) showed extremely high levels of hydrocarbons. Elevated levels were still noted in mussel tissues at Alyeska Marine Terminal during cruise 4 (July 1994), and a strong signal was also observed in the Gold Creek samples at this time.

A similar spike in the mussel contamination from a weathered ANS source was noted in the samples collected at Disk Island during cruise 4 (July 1994), and at first it might seem plausible to speculate that it too was from the *Eastern Lion* oil spill. However, sheens released from mussel bed cleaning operations at Disk Island the day before the RCAC samples were collected (Jeff Short, personal communication) represent a far more likely source.

After the July 1994 time frame, hydrocarbon levels in the mussels in these and all other stations dropped to uniformly low values in cruise 6 (July 1995). The PAH histogram pattern for the mussels collected from Gold Creek (Figure 4-14, page 71) is indicative of the extremely low background signal observed in mussel samples throughout Prince William Sound at that time. Based on the diagnostic ratios of individual constituents in this histogram, it is identified as being characteristic of the by-products associated with the combustion of oil. It should be noted, however, that this same pattern was sometimes reported in field and laboratory blanks, although at concentrations that were generally a factor of 10 lower. Therefore, although

this pattern may be considered by some to be an artifact of the sampling and measurement program, it does show up consistently at low-level sites in other monitoring efforts such as the NOAA Status and Trends program. As such, it may actually reflect the low-level hydrocarbon signals associated with combustion products from burning diesel and/or fuel oil in the vicinity of the site.

Examination of the *Mytilus* Petrogenic Index plot shows another increase at all stations during cruises 8 and 9 (July 1996 and March 1997). The profiles obtained in these samples are again consistent with those observed for ANS crude oil at Alyeska Marine Terminal, Gold Creek, and Disk Island. The cruise 9 profiles for the increase observed at Sleepy Bay and Knowles Head, however, are not consistent with the source being ANS crude oil. One possible source for a newly arising signal observed in the last sampling interval (at least for Alyeska Marine Terminal and Gold Creek), would be from the Alyeska Ballast Water Treatment Plant (BWTP) spill, which occurred in January 1997.

The histogram profiles associated with the increases in hydrocarbon concentrations at Sleepy Bay during cruises 3, 4, and 5 show that the source is consistent with EVOS or more recent releases of ANS crude oil.

The interpretation of the *Mytilus* Petrogenic Index pattern at Windy Bay is somewhat more complicated. A mixed source is indicated, including contributions from aromatics that look like they could be derived from Bunker C or No. 6 fuel oil. In addition, the contributions from biogenic hydrocarbons (plant waxes and natural oils), as measured in the sediments, are higher at Windy Bay than at any other site. This site also contains traces of ANS or EVOS-related oil that were detected in the deep sediments during cruises 3, 4, 6, and 8. However, hydrocarbons associated with EVOS oil were not observed in the shallow sediments. ANS or EVOS oil was only observed in the mussel samples collected during cruise 9 at Windy Bay. The Kinnetics field samplers suggest that logging operations in the area may be a new possible source of both biogenic and petrogenic hydrocarbons.

At this phase of the program, several facts seem well established.

1. The deep sediments serve as good long-term integrators of hydrocarbon input to the region, but for the most part, they do not reflect recent events or temporal impacts from known point source releases in the recent past.
2. The temporal variability of PAH and AHC levels in sediments is low except at Alyeska Marine Terminal and Sleepy Bay.
3. Absolute PAH levels in the shallow sediments are approximately an order of magnitude lower, and they do not always show the same compositional pattern, as their deeper sediment counterparts.
4. The mussel tissue sampling and analysis program is achieving its goal of detecting low-level changes in hydrocarbon burdens and in identifying the presence or absence of Alaskan North Slope crude oil.
5. The analytical variability associated with triplicate measurements of mussel tissue is small enough to detect differences among sites within the goals of the program. Temporal trends within sites can be detected with a moderate degree of confidence, but they are currently constrained by the short period of sampling.
6. Hot spots for mussel tissues are areas where the *Mytilus* Petrogenic Index exceeds 400 (Table 5-1).
7. Even though there are some potential problems with field and procedural blanks, the patterns obtained for low-concentration histograms will not interfere with the detection of ANS or EVOS oil at extremely low levels.
8. Overall, the signal for EVOS oil is low and sporadic, suggesting that either not much EVOS oil remains in most locations in PWS or that at least it is relatively inaccessible.
9. There are chronic inputs of hydrocarbons to the marine system, but residual levels are very low.

6. Recommendations for Modifications to LTEMP

Table 6-1 summarizes our general assessments of the monitoring program.

Table 6-1. RCAC Monitoring Program Assessments

Sample Type	Program Assessment
Deep sediments	Working well, sampling is consistent and detection levels are good, but samples are only detecting background hydrocarbons (i.e., non-ANS).
Shallow sediments	Generally working well, but there are problems with detection limits and sampling depths.
Mussel tissues	Hydrocarbon detection is working very well, but the morphometric and lipid data being collected are unnecessary.

6.1 Strategy

Overall, the program is working well although the station coverage is a somewhat sparse. The use of mussels as a sentinel organism within Prince William Sound is successful. Although the measured hydrocarbon levels are low, and there are potential problems with interferences due to background contaminants associated with field and laboratory procedures (particularly at very clean sites), these low-level concentrations will allow a very minor increase in hydrocarbon concentrations to be detected. While such detection is difficult to do on a statistically significant basis, the utilization of characteristic patterns (which involve individual compounds such as those shown in the histogram plots presented in this report) makes it easier to identify trends and changes in sources, even when absolute total hydrocarbon loadings may not be changing that much. As such, the overall TPAH or TAHC concentration value at a site may not change that much, but the influence of a new source can be readily identified by a change in the histogram pattern for constituent compounds.

We were able to assess sources, track patterns, correlate with events, and detect a few statistically significant differences within the existing data; but there were definite constraints due to small sample size (i.e., number of sites and sampling intervals). It may be possible to monitor larger reaches of the sound on a limited budget by reducing the sampling efforts, for example, by sampling during one season rather than two or by changing to biennial samplings at "stable" sites. If, under this looser but broader-focus sampling, an acute change were detected or a catastrophic event occurred, the program could still respond with increased intensity of monitoring in an affected locale.

6.2 Sample Types

Several options are available for restructuring the monitoring program. Table 6-2 assesses the pros and cons of continuing the three types of samples currently being collected and addresses the option of adding intertidal sediments, an alternative that has been suggested.

Table 6-2. Assessment of Current and Potential Sample Types

Sample Type	Advantages	Disadvantages
Deep sediments	The depositional sediment regime is appropriate for assessing historic fluctuations in background hydrocarbons.	The hydrocarbons measured are almost exclusively background from seeps or coal transported into the sound. There is little correlation between deep and shallow sediment hydrocarbons. Thus, it is unlikely that a surface spill will be detected in significant amounts in deep sediments.
	The hydrocarbon levels are more stable and typically exceed those in shallow sediments.	Subtidal sediments do not acquire the dissolved water-soluble fractions from oil as mussels do.
		Relative to sampling the shallow sediments, the risk of sampling failure due to inclement weather, equipment malfunction, sample handling, or station keeping is higher.
Shallow sediments	Hydrocarbons from intertidal spills are transported relatively rapidly to the shallow subtidal sediments.	Shallow sediments are exposed to a higher energy regime relative to deeper sediments. Thus, the coarser sediment matrix rapidly loses its hydrocarbon loads either through dispersion or weathering.
	Diver sampling is tedious and at risk of being weathered out but less liable to failure than deep sediment grab sampling.	Subtidal sediments do not acquire the dissolved water-soluble fractions from oil as mussels do.
		There is little apparent correlation between shallow subtidal sediment hydrocarbons and mussel tissue loads (and presumably, intertidal sediments).

Table 6-2 continues

Table 6-2 (continued**)**

Intertidal sediments	Intertidal sediments receive the bulk of deposited hydrocarbons following an oil spill. All forms of hydrocarbons (e.g., gross contamination by fresh crude containing water-soluble fractions and heavier constituents) may be represented in intertidal sediments.	Hydrocarbons are patchily distributed both horizontally and vertically within the sediments and along the intertidal slope. This is a function of beach exposure, sediment type, and the chance involved in grounding of wave-deposited hydrocarbons. It is difficult to select a single sampling location that is representative of a contaminated beach.
	Mussel tissue loads correlate better with intertidal than subtidal sediments.	
	The risk of sampling failure is much lower than for subtidal sediments. Identification of sampling stations is straightforward, and accurate resampling is facilitated by the use of landmarks, flagged markers (rebar), or other prominent geographical features.	
Mussel tissues	Mussels are very sensitive indicators of contamination by both water-soluble fractions and discrete oil droplets (very fine particulate hydrocarbons).	Depuration and metabolism create a limited temporal window for detection of low-level events.
	The risk of sampling failure is much lower than for subtidal sediments.	Body loads vary with body size. Sampling consistency is paramount.

6.3 Statistical Methods

For this type of study, we recommend loosening the alpha level from 0.05 to 0.10 and tightening the beta (currently 0.20 with the target power of 0.8). Loosening the alpha would improve the power to detect a difference at lower levels (currently not a problem at the 25-percent level for most sites and variables) and reduce the odds of mistakenly rejecting a sample that may be significantly different. These goals mesh better with RCAC objectives of monitoring for occurrence of oil in the environment than the classically more conservative alpha which sets a higher goal for a difference to be declared significant.

6.4 Sampling Methodology

- Shallow sediment collections are a little sloppy relative to depth (both within and between cruises). Be more careful about resampling at the same depth, particularly at Sheep Bay, Disk Island, and Sleepy Bay.
- The depth dependence phenomenon appears to be more of a problem with the shallow sediments. However, at Sheep Bay there may also be a depth dependence noted for the deeper sediments collected in cruises 6 and 8. Alternatively, we might have identified a temporal trend.
- If mid-depth sediment sampling is continued, increase the size of sample extracted to lower the detection limits. Most measurements for PAH and TAHC in the shallow sediments are at or just below detection limits.
- Analyses of field blanks and equipment rinses indicate some procedural problems. Hydrocarbon analyses of the field blanks and many of the lower level mussel and shallow sediment samples look very similar, if not identical.
- Tighten up the size specifications for mussels at all sites to reduce the potential for size-related differences in PAH.

6.5 Analytical Techniques

- Analyze diesel soot and determine if it contributes to the pattern observed in the field blanks and the low-level samples. Bence and Burns (1995) refer to this pattern as “Procedural Artifacts.” It certainly

shows up in a majority of analyses for the field blanks and equipment rinses. NOAA (1997) has reported on the PAH pattern obtained from by-products of oil combustion; however, no data were readily available for diesel soot.

- Have GERG rerun batches if a certain number of analytes in blanks are present above their individual method detection limits. Current standard operating procedures identify procedural blanks as a problem only if individual analytes are 3x MDL. However, the absolute number of analytes that can be >3x MDL is not specified. Many of the procedural blanks contain PAHs, which although not 3x the individual MDLs are at levels that approach the concentrations in the mussels and shallow sediments from the cleaner stations.
- Discontinue analysis for lipids for the purpose of making lipid correction (reporting PAH on basis of ng/g lipid). Lipid correction appeared to distort the temporal patterns, eliminating measurable spikes or increases in uncorrected hydrocarbon concentrations at the Alyeska Marine Terminal even when known spill events occurred in the immediate vicinity of the sampling stations. Most other biological measurements (shell volume and measures of reproductive state (gonad vs. somatic tissue) should be discontinued. None of those measurements show any relationship to anything. Nevertheless, it is important to continue measuring shell length to provide confirmation that the same general size range (and therefore approximate age) of mussels was collected at all sites during all sampling events.
- Analyze other potential sources of PAH to sediments (and potentially biota) including samples of Katalla seep oil and coal samples from the deposits to the east of Prince William Sound. In this regard, it will be critically important to ensure that the characterization data are obtained by the same laboratory and analytical procedures currently used for the LTEMP mussel tissue and sediment samples.
- Reinstate the analyses of aliphatic hydrocarbons in mussel tissues. Require GERG to implement a cleanup procedure (e.g., SiO_2 chromatography or SepPacs®) that will remove interfering lipids to allow accurate quantification of n-alkanes and other branched and cyclic aliphatics, if possible. Also, in this regard, request that they report the total resolved aliphatic hydrocarbons, not just the n-alkanes plus pristane and phytane. Aliphatic hydrocarbon measurements are important for differentiating the relative contributions of biogenic hydrocarbons from other (petrogenic and pyrogenic) sources. Even if MDLs for the aliphatic hydrocarbons are significantly higher that those for PAH, these measurements will be valuable because: 1) biogenic alkanes are the most predominant hydrocarbons currently being introduced to the Prince William Sound ecosystem and this should be reflected in your database; and 2) in almost all crude oils and refined products, the aliphatic hydrocarbons are anywhere from 10 to 100 times more abundant than the PAH. As such, they will be easily detected in mussel tissue samples (even with higher MDLs) in the event of a spill, and they can aid in discriminating among numerous potential sources (crude oil vs. distilled products vs. lube oils, etc.).

6.6 Sampling Seasons

Based on the analyses of mussel and sediment data, it is conceivable that the seasonal samples could be reduced to one season. PAH concentrations appear to be higher for mussel samples in winter so this should improve the detection limit issues. Moreover, spikes were more common in winter samples. However, variability is marginally lower in summer (coefficients of variation were 20.3 and 26.8 percent in summer and winter, respectively) and the risk and cost associated with sampling are lower in summer. If cost reductions are obtained by eliminating an entire sampling season, it may be possible to add several stations to increase sampling coverage.

6.7 Sampling Locations

If RCAC is able to cut back on seasonal samplings, we would recommend adding sites to expand the area of coverage and monitoring hydrocarbon exposure from human activities and background sources.

- Based on the suggestions and evidence from Page et al. (1997) and Short and Heintz (1997), the deep stations are continuously exposed to background hydrocarbons transported either from oil seeps or coal deposits outside of PWS. We saw very little evidence of significant quantities of EVOS/ANS residues in any at the deep stations; shallow subtidal oil apparently becomes widely dispersed or highly weathered once it leaves the shallow depths. If sampling at deep stations is continued, we recommend that a new sampling site be established in Hinchinbrook Entrance, either in Constantine Harbor, the entrance to Port Etches, or Zaikof Bay, to provide insight into the signatures, concentrations, and flux

of the "background" hydrocarbons entering Prince William Sound. Alternatively, if deep sediment sampling is dropped, we would recommend at least a special sampling of sediments and potential sources (Katalla crude and coal) from outside PWS to benchmark the signature of the transported sources.

- We recommend that the station network be expanded to other regions of the Sound, e.g., the eastern, north-central, and northwestern Sound. In the event that another catastrophic spill like EVOS causes exposure to ANS crude in other regions in the Sound beyond the monitoring area, little information regarding current hydrocarbon loading in sediments or mussel tissues exists.
- All hydrocarbon studies conducted in the Sound find miscellaneous residues of petroleum hydrocarbons, typically from diesel fuels, bunker fuel, and combustion products. In 1991, Houghton et al. (1992) observed substantially higher concentrations of PAHs in mussel tissue from the vicinity of the cannery in Seward than from sites in Prince William Sound that previously had been grossly contaminated by EVOS. In 1995, Roberts et al. (1996) noted that the highest PAH levels from their limited sampling of mussels were found in the harbor at Whittier. It seems obvious to expect inputs from the currently unmonitored human activities near towns and villages, marinas, hatcheries, ferry docks, airports, and logging operations throughout the Sound. A few screening samples from areas near human activities would help assess the need for additional sites and provide helpful insights.

7. References Cited

Arkhipova, E.A., and N.V. Shadrin. 1992. Gonad cycle of trhe Pacific mussel *Mytilus trossulus* from Avacha Inlet, Eastern Kamchatka. Biologiya Morya - Marine Biology 34(1-2):96-100.

Bence, A.E., and W.A. Burns. 1995. Fingerprinting hydrocarbons in the biological resources of the *Exxon Valdez* spill area. In: P.G. Wells, J.N. Butler, and J.S. Hughes (Eds.), *Exxon Valdez* Oil Spill: Fate and Effects in Alaskan Waters. American Society for Testing and Materials (ASTM), Philadelphia, PA. pp 84-140.

Boehm, P.D., and J.W. Farrington. 1984. Aspects of the polycyclic aromatic hydrocarbon geochemistry of recent sediments in the Georges Bank Region. Environmental Science and Technology 18:840-845.

Boehm, P.D., and A.G. Requejo. 1988. Overview of the recent sediment hydrocarbon geochemistry of Atlantic and Gulf Coast over continental shelf environments. Estuarine, Coastal and Shelf Science 23:29-58.

Brassell, S.C., G. Eglinton, J.R. Maxwell, and R.P. Philip. 1978. Natural background of alkanes in the aquatic environment. In: 0. Huntzinger, L.H. van Lelyveld, and B.C.J. Zoetman (Eds.), Aquatic Pollutants, Transformations and Biological Effects. Pergamon Press, Oxford. pp 69-86.

Dare, P.J., and D.B. Edwards. 1975. Seasonal changes in flesh weight and biochemical composition of mussels (*Mytilus edulis* L.) in the Conwy Estuary, North Wales. J. Exp. Mar. Biol. Ecol. 18:89-97.

Driskell, W.B., J.R. Ruesink, D.C. Lees, J.P. Houghton, R. Gilmour, S.C. Lindstrom, A. Mearns, in prep. Seven-year signal of intertidal disturbance following the *Exxon Valdez* Oil Spill.

Edgington, E.S. 1987. Randomization Tests. New York, Marcel Dekker. 341 pp.

Farrington, J.W., and B. W.Tripp. 1977. Hydrocarbons in western North Atlantic surface sediments. Geochim. Cosmnchim. Acta 41:1627-1641.

Feder, H.M., and D.G. Shaw. 1987. Environmental studies in Port Valdez, Alaska, 1987. Final Report to Alyeska Pipeline Service Company. Institute of Marine Science, University of Alaska, Fairbanks. 175 pp.

Feder, H.M., and D.G. Shaw. 1990. Environmental studies in Port Valdez, Alaska, 1989. Final Report to Alyeska Pipeline Service Company. Institute of Marine Science, University of Alaska, Fairbanks. 297 pp.

Federal Register. 1986. Code of Federal Regulations, Title 40, Part 136, Protection of the Environment. (40 CFR 136). U.S. Government Printing Office, Washington, D.C.

Gosling, E. (Ed.) 1992. The Mussel, *Mytilus*: Ecological, Physiological, Genetics, and Culture. Vol. 25. Developments in Aquaculture and Fisheries Science. Elsevier. New York. 589 pp.

Houghton, J.P., A.K. Fukuyama, D. C. Lees, H. J. Teas II, T. A. Ebert, and W. B. Driskell. 1992. Evaluation of the 1991 condition of Prince William Sound littoral biota following the *Exxon Valdez* oil spill and subsequent shoreline treatment. For National Oceanic and Atmospheric Administration, Ocean Assessment Division.

Karinen, J.F. M.M. Babcock, D. W. Brown, W. D. MacLeod, Jr., L. S. Ramos, and J., W. Short. 1993. Hydrocarbons in inertidal sediments and mussels from Prince William Sound, Alaska, 1977-1980: Characterization and probable sources. NOAA Tech Memorandum NMFS-AFSC-9. U.S. Dept. of Commerce, National Marine Fisheries Service.

Kernicut, M.C., II, and P. Comet. 1992. Resolution of sediment hydrocarbon sources: multiparameter approaches. In: J.K. Whelan and J.W. Farrington (Eds.), Organic Productivity, Accumulation, and Preservation in Recent and Ancient Sediments. Columbia University Press. pp 308-337.

Kimball, D.M., and A.E. McElroy. 1993. Characterizing the annual reproductive cycle of *Mytilus edulis* from Boston Harbor and Cape Cod Bay - a comparison by means of stereology and condition indices. Marine Environmental Research 35(1-2):189-196.

Kinnetic Laboratories, Inc. 1993. Power analysis report and recommendations for sampling based on subset of highly replicated data from LTEMP survey number one. Prepared for Scientific Advisory Committee of Prince William Sound RCAC.

Kinnetic Laboratories, Inc. 1997. Prince William Sound RCAC Long-Term Environmental Monitoring Program. Publ. No. C/608.961\LTEMP. November 1997.

Kvenvolden, K.A., F.D. Hostettler, P.R. Carlson, and J.B Rapp. 1995. Ubiquitous tar balls with a California-source signature on the shorelines of Prince William Sound. Alaska. Env. Sci. & Tech. 29 (10):2684- 2694.

National Oceanic and Atmospheric Administration. 1989. A summary of data on tissue contamination from the first three years (1986-1988) of the Mussel Watch Project. NOAA Technical Memorandum NOS OMA 49. U. S. Dept. of Commerce, National Oceanic and Atmospheric Administration, Rockville, MD. 22 pp.

National Oceanic and Atmospheric Administration. 1993. Sampling and analytical methods of the National Status and Trends Program National Benthic Surveillance and Mussel Watch Projects 1984-1992. Volume IV. Comprehensive description of trace organic analytical methods. NOAA Technical Memorandum NOS ORCA 71.

National Oceanic & Atmospheric Administration. 1997. Integrating Physical and Biological Studies of Recovery from the *Exxon Valdez* Oil Spill; Case Studies of Four Sites in Prince William Sound, 1989-94. Gary Shigenaka, ed. NOAA Technical Memorandum NOS ORCA 114. 206 pp + appendix.

O'Connor, T.P., A.Y. Cantillo, and G.G. Lauenstein. 1994. Monitoring of temporal trends in chemical contamination by the NOAA National Status and Trends Mussel Watch Project. IN: K. J. M. Kramer, ed., Biomonitoring of Coastal Waters and Estuaries, CRC Press, Boca Raton, FL. pp 29-50.

Page, D.S., P.D. Boehm, G.S. Douglas, and A.E. Bence. 1995. Identification of hydrocarbon sources in the benthic sediments of Prince William Sound and the Gulf of Alaska following the *Exxon Valdez* oil spill. In: P.G. Wells, J.N. Butler, and J.S. Hughes (Eds.), *Exxon Valdez* Oil Spill. Fate and Effects in Alaskan Waters. American Society for Testing and Materials (ASTM), Philadelphia, PA. pp 41-83.

Page, D.S., P.D. Boehm, G.S. Douglas, A.E. Bence, W.A. Burns, and P.J. Mankiewicz. 1996. The natural petroleum hydrocarbon background in subtidal sediments of Prince William Sound, Alaska, USA. Environ. Tox. Chem. 15(8):1266-1281.

Page, D.S., P.D. Boehm, G.S. Douglas, A.E. Bence, W.A. Burns, and P.J. Mankiewicz. 1997. An estimate of the annual input of natural petroleum hydrocarbons to seafloor sediments in Prince William Sound, Alaska. Mar. Poll. Bull. 34(9):744-749.

Payne, J.R., J.R. Clayton, Jr., G.D. McNabb, Jr., and B.E. Kirstein. 1991a. *Exxon Valdez* oil weathering fate and behavior: Model predictions and field observations. In: Proceedings of the 1991 Oil Spill Conference, American Petroleum Institute, Washington, D.C., pp 641-654.

Payne, J.R., J.R. Clayton, Jr., C.R. Phillips, J.L. Lambach, and G.H. Farmer. 1985. Marine oil pollution index. Oil and Petrochem. Poll. 2(3):173-191.

Payne, J.R., L.E. Hachmeister, G.D. McNabb, Jr., H.E. Sharpe, G.S. Smith, and C.A. Manen. 1991b. Brine-induced advection of dissolved aromatic hydrocarbons to arctic bottom waters. Environmental Science and Technology 25(5):940-951.

Payne, J.R., B.E. Kirstein, G.D. McNabb, Jr., J.L. Lambach, C. de Oliveira, R.E. Jordan and W. Hom. 1983. Multivariate analysis of petroleum hydrocarbon weathering in the subarctic marine environment. In: Proceedings of the 1983 Oil Spill Conference. American Petroleum Institute, Washington, D.C., pp 423-434.

Payne, J.R., B.E. Kirstein, G.D. McNabb, Jr., J.L . Lambach, R. Redding, R.E. Jordan, W. Hom, C. de Oliveira, G.S. Smith, D.M. Baxter, and R. Geagel. 1984. Multivariate analysis of petroleum weathering in the marine environment - subarctic. Volume I, Technical Results; Volume II, Appendices. In: Final Reports of Principal Investigators, Vols. 21 and 22. February 1984, U.S. Department of Commerce, National Oceanic and Atmospheric Administration, Ocean Assessment Division, Juneau, Alaska. 690 pp.

Payne, J.R., and G.D. McNabb, Jr. 1984. Weathering of petroleum in the marine environment. Marine Technology Society Journal 18(3) 24-42.

Payne, J.R., G.D. McNabb, Jr., and J.R. Clayton, Jr. 1991c. Oil-weathering behavior in arctic environments. Polar Research 10(2):631-662.

Pruell, R.J., J.G. Quinn, J.L. Lake, and W.R. Davis. 1987. Availability of PCBs and PAHs to *Mytilus edulis* from artificially resuspended sediments. In: J.M. Capuzzo and D.R. Kester (Eds.), Oceanic Processes in Marine Pollution. Vol. 1. Biological Processes and Wastes in the Ocean. Kreiger, Malabar, FL. pp 97-108.

Roberts, P., C.B. Henry, E.B. Overton. 1996. Evaluation of the condition of Prince William Sound shorelines following the *Exxon Valdez* oil spill and subsequent shoreline treatment. NOAA Technical Memorandum NOS ORCA 91, Vol. III, 1994 Summary of Chemistry Results. 67 pp + appendices.

Sericano, J.L., E.L. Atlas, T.L. Wade, and J.M. Brooks. 1990. NOAA's Status and Trends Mussel Watch Program: Chlorinated pesticides and PCBs in oysters (*Crassostrea virginica*) and sediments from the Gulf of Mexico 1986-1987. Mar. Environ. Res. 29:161-203.

Shaw, D.G., B.B. Bergeron, C.R Geist, D.L Hawkins, and D.J. McIntosh. 1993. Hydrocarbon studies in Port Valdez, Alaska: 1992. In: H.M. Feder and D.G. Shaw (Principal Investigators), Environmental Studies in Port Valdez, Alaska: 1992. Final Report. Instiutute of Marine Science, University of Alaska, Fairbanks, AK. Submitted to Alyeska Pipeline Service Company Anchorage, AK. pp 5-134.

Shaw, D.G., B.E. Bergeron, D.L Hawkins, C.R Geist, and D.J. McIntosh. 1994. Hydrocarbon studies in Port Valdez, Alaska: 1993. In: H.M. Feder and D.G. Shaw (Principal Investigators), Environmental Studies in Port Valdez, Alaska: 1993. Final Report. Instiutute of Marine Science, University of Alaska, Fairbanks, AK. Submitted to Alyeska Pipeline Service Company Anchorage, AK. pp 11-137.

Shaw, D.G., T.E. Hogan, and D.J. Macintosh. 1985. Hydrocarbons in the sediments of Port Valdez, Alaska: consequences of five years' permitted discharge. Estuarine, Coastal and Shelf Science 21:131-144.

Shaw, D.G., T.E. Hogan, and D.J. Macintosh. 1986. Hydrocarbons in bivalve mollusks of Port Valdez, Alaska; consequences of five years' permitted discharge. Estuarine, Coastal and Shelf Science 23:863-872.

Short, J.W., and M.M. Babcock. 1996. Prespill and postspill concentrations of hydrocarbons in mussels and sediments in Prince William Sound. In: S.D. Rice, R.B. Spies, D. A. Wolfe, and B.A. Wright, Eds. Proceeding of the *Exxon Valdez* oil spill symposium. American Fisheries Society Symposium 18. pp 149-166.

Short, J.W., and R.A. Heintz. 1997. Identification of *Exxon Valdez* oil in sediments and tissues from Prince William Sound and the Northwestern Gulf of Alaska based on a PAH weathering model. Environmental Science & Technology, 31(8):2375-2384.

Thompson, R.J. 1984. The reproductive cycle and physiological ecology of the mussel *Mytilus edulis* in a sub-arctic, non-estuarine environment. Mar. Biol. 79:277-288.

Venkatesan, M.I. 1988. Occurrence and possible sources of perylene in marine sediments – a review. Marine Chemistry, 25:1-27.

Widdows, J., K.A. Burns, N.R. Menon, D.S. Page, and S.Soria. 1990. Measurement of physiological energetics (scope for growth) and chemical contamination in

mussels (*Arca zebra*) transplanted along a contamination gradient in Bermuda. J. exp. mar. Biol. Ecol. 138:99-117.

Appendix VIII. Identification of oil source by applying Short and Heintz oil weathering model to LTEMP sediment and Mytilus tissue PAH data

Concept of Model

The Short and Heintz study found that oil weathering could be modeled as simple first order mathematical equations for each of 14 PAH compounds typical of ANS/EVOS crude oil. Thus, given a hydrocarbon sample from Prince William Sound, they could determine how weathered a sample was based on the relative abundance of each of the remaining PAHs. Then they would compare the sample with the expected values predicted for EVOS oil degraded to the same weathering state. The statistical result of the comparison, the mean square difference (MSE), is simply the fit or difference between the two samples. For example, how much does the histogram bar for each (log transformed) PAH protrude above the same one in the comparison sample? Square the difference, sum for all PAHs and divide by the number of PAHs compared. A larger MSE implies a less perfect fit.

Application of Model to RCAC Data

In the following appendix, Short calculated the MSE for fitting the sample to both an EVOS standard sample and a sample of the "background" PAH signal presumably from the Katalla oil seeps or coal deposits. The background sample was actually taken from a non-impacted site, Constantine Harbor at Hinchenbrook entrance, the site of the dominant tidal influx into Prince William Sound. The fit to the background sample is labeled as MSE'. Short then uses a probability model to predict threshold values for identifying the sample as being derived from EVOS or Constantine source oil (labeled as E or C in the table). If the MSE result is marginally beyond the probability threshold but obviously closer to one source than another, then subjectively, a parenthetical (E) or (C) is assigned.

In the original model, Short and Heintz set acceptance criteria for sample validation prior to applying the model. They required that the sample's TPAH exceed MDL by 100 ng/g to produce credible results, and also preferred the TPAH be larger than 750 ng/g for good source discrimination. For the RCAC data set, the criteria were dropped because they would have eliminated nearly all of the data set. Normally, each of the three sample replicates would appear in the appendix; however, replicates that could not be analyzed appear as data gaps. These samples had zero levels for some PAHs, which meant the model, could not be applied.

The original model was published in J. W. Short and R. A. Heintz, 1997. Identification *of Exxon Valdez* oil in sediments and tissues from Prince William Sound and the Northwestern Gulf of Alaska based on a PAH weathering model. Environmental Science & Technology, 31:8, 2375-2384.

Appendix Notations:
MSE = Mean square error differences between measured PAHs and PAH levels predicted for EVOS at a state of weathering matching that of the sample.

MSE' = MSE differences between measured and predicted Constantine PAH

Source ID Criteria:

C = Constantine-type source (natural background) if MSE' < 0.34

E = EVO source if MSE < 0.98

Marginal Cases (subjectively assigned):

(C) = Much closer to the Natural Background than to EVO

(E) = Much closer to EVO than the Natural Background

Email comments by Jeff Short on the data sets and the results of his analyses:

Sediment Samples

Following the initial run using original sample rejection criteria:

"I was surprised that only 35 samples met the MDL criteria given for valid model application, especially given the high TPAH values of some of the excluded samples. The MDLs for these samples were surprisingly high - 20 ng/g or more....

Among the samples that qualified for model application, I was generally pleased that the results conformed to my a priori expectations. The higher TPAH samples at AMT correspond with very weathered ANS, as would be expected in the subtidal sediments deeper than a few meters there. The deeper sediments elsewhere generally correspond with the pattern characteristic of the natural background. The model is pretty sensitive to pattern deviations, so I have indicated cases where it is closer to a suspected source by use of (). Note that the 2 high TPAH samples from Disk Island intertidal are dead ringers for not very weathered EVO."

After relaxing sample rejection criteria:

"I found the [RCAC LTEMP sediment PAH data] results conform completely to my expectations. Note that one of the reasons that a sample might not fit either modeled source is if it's a mixture of both. When I mentioned the 750 ng/g [TPAH] cutoff in the paper, my point was that when samples contain more than this, they always fell clearly into one of the two modeled sources. Below this, classifications of neither (i.e. a possible mixture of the two) or both (i.e., indeterminate) sometimes occur (but not necessarily). The picture I see in this data set confirms that the bottom of PWS is covered by the natural background source, as are deeper sediments sampled from the GOA, and that ANS (possibly mixed with the natural background) occurs at AMT and intermittently at Sleepy Bay and Disk Island, which is not a surprise."

Mussel Tissues

"Unfortunately, only 19 [of the mussel data set] samples had non-zero entries for the modeled analytes. This is 'very' surprising, given the high TPAH levels reported....

Of the 19 samples considered, all but three are dead ringers for ANS. Two of the remainders are close, and all 3 of the outliers are the result of anomalous C2- & C3 fluorine concentrations, which is probably the result of an analytical interference at GERG (we have occasionally had similar problems here too). These analytes are especially susceptible to interfering ions from e.g.

the surrogates, and as I mentioned earlier it doesn't take much to perturb the model. Note that the *Eastern Lion* samples are a dead match for not very weathered ANS.”