Exxon Valdez Oil Spill
Restoration Project Final Report

Evaluation of Oil Remaining in Prince William Sound from the *Exxon Valdez* Oil Spill

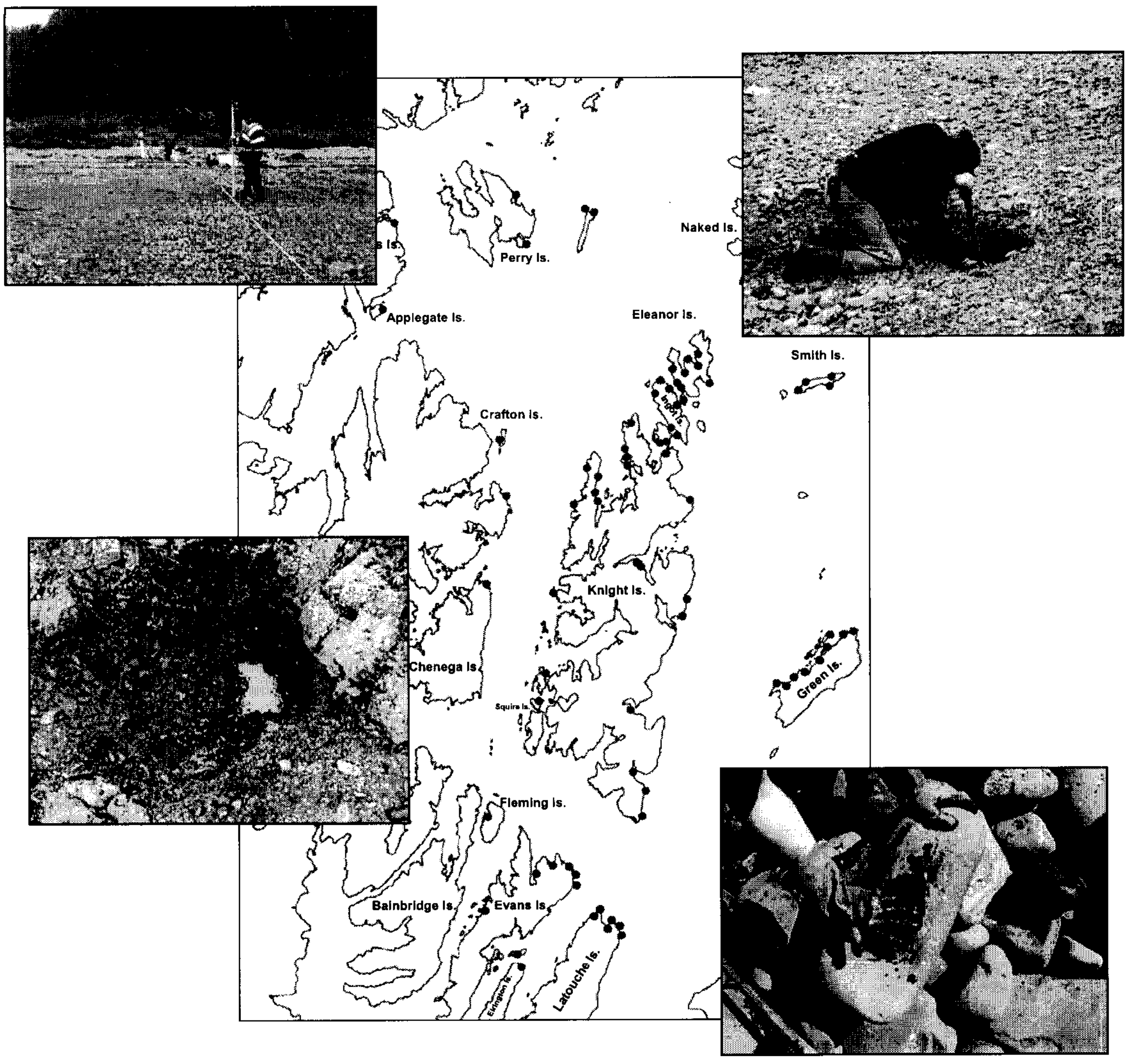
Restoration Project 02543
Final Report

Jeffery W. Short
Mandy R. Lindeberg
Patricia M. Harris
Jacek M. Maselko
Jerome J. Pella
Stanley D. Rice

Auke Bay Laboratory, Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA

April 2003

Evaluation of Oil Remainaing in Prince William Sound from the *Exxon Valdez* Oil Spill



By
Jeffery W. Short
Mandy R. Lindeberg, Patricia M. Harris,
Jacek M. Maselko, Jerome J. Pella, and Stanley D. Rice

Auke Bay Laboratory, Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA.

Exxon Valdez Oil Spill
Restoration Project Final Report

Evaluation of Oil Remaining in Prince William Sound from the *Exxon Valdez* Oil Spill

Restoration Project 02543
Final Report

Jeffery W. Short
Mandy R. Lindeberg
Patricia M. Harris
Jacek M. Maselko
Jerome J. Pella
Stanley D. Rice

Auke Bay Laboratory, Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA

April 2003

Short, J., M.R. Lindeberg, P.M. Hatrris, J.M. Maselko, J.J. Pella, and S.D. Rice. 2003. Evaluation of Oil Remaining in Prince William Sound from the *Exxon Valdez* Oil Spill. EVOSTC Restoration Project 02543 Final Report.

Table of Contents

Executive Summary

Reasons for the study

This project was undertaken to quantitatively evaluate the amount of oil from the *Exxon Valdez* oil spill (EVOS) that remained in Prince William Sound (PWS), Alaska, during summer 2001. Following the ten year anniversary in 1999, pockets of oil were still being found by the public, scientists, and the media. The quantity and significance of this oil was unknown, stimulating the *Exxon Valdez* Trustee Council (EVTC) to fund a study that would perform the first evaluation of remaining since 1993. This report presents the results and findings of this study. In the summer of 2001, over 90 sites were evaluated for remaining deposits of surface and sub-surface oil, and subtidal oil at 5 transects sampled in 1989-91.

This project has three goals: (1) to determine the amount of oil and oiled shoreline remaining in PWS, (2) to determine the recovery rates of oiled shorelines and of subtidal sediments, and (3) to verify that the oil remaining was *Exxon Valdez* oil (EVO) by chemical fingerprinting. The first task was the most challenging. Initially it seemed logical to simply repeat the methods used during the years immediately following the EVOS so that the results would be directly comparable, but this approach had serious implicit shortcomings. The method used to evaluate oil persistence immediately after the EVOS relied on shoreline cleanup assessment teams (SCAT), who comprehensively inspect beaches visually and record results following standardized procedures, hereafter termed the SCAT method. Because oil may become incorporated into subsurface sediments, these visual assessments were augmented by excavations where oil was seen to be most persistent. This procedure was adequate for directing clean-up operations, but was poorly equipped to provide quantitative estimates of oil remaining or area of beach contaminated. After twelve years, the visual SCAT procedures may additional inadequacies: remaining surface oil may no longer provide reliable visual cues for locating subsurface oil. Lastly, the available resources were insufficient to repeat a comprehensive survey of the entire impacted area within PWS. Procedures used in 2001 would have to cover a sub-set of the total number of beaches impacted in 1989 (fewer resources available), considerable effort would be needed to evaluate subsurface oil, quantitative procedures would be needed, and a measure of the precision (statistical confidence interval) of the estimate would be required. The random sampling approach that was ultimately adopted substantially addressed these issues.

Methods

Random sampling of oiled shoreline

The random sampling method used presented its own challenges, because such an approach had never before been applied to the problem of estimating the amount of oil on a beach, which engendered concerns regarding its feasibility. Because there was no reason to discount the claims of Exxon's contractors that there was very little oil present in 2001, the sampling design had to consider the possibility that the oil encounter rate during a random sampling of beach sediments might be very low, and conceivably zero. A random sampling that failed to encounter oil would lead to a trivial result at considerable expense, so the sampling was carefully designed to minimize this possibility. The sampling was stratified, or focused on the most heavily oiled beaches, where the oil was most likely to persist. Because oil was incorrectly thought to be mostly in the upper intertidal, the sampling was limited to the upper half of the intertidal. The particular design used, called adaptive sampling, is especially appropriate for sampling something that is rare and highly aggregated. Adaptive sampling is an extension of random sampling, where once a patch of oil is located by random sampling, the size of the patch is then estimated by very, localized additional sampling. It turned out that this statistical refinement was unnecessary, because oil was encountered so frequently that random sampling alone sufficed to produce statistically valid estimates. But crucial to both sampling designs was the random selection of a large number of beaches, and excavation of a large number of pits on each beach to search for subsurface oil. We selected 91 beaches for sampling, and dug a total of nearly 9,000 test pits (including 6,775 randomly drawn pits) over the summer of 2001. Results from both the random sampling alone, and from the adaptive sampling, are presented as separate chapters in this report. These are the first fully quantitative estimates of beached oil for any oil spill anywhere, and hence provide a groundbreaking advance in the determination of oil persistence on beaches.

Trends over time and subtidal evaluations

Trends in the recovery of inter- and subtidal sediments were determined by comparisons with previous surveys. For the intertidal, the results of the random sampling effort could be compared with results from the SCAT surveys of 1989 - 1993, but these comparisons cannot be direct because of the differences in methods used. To address this, one element of this project was devoted to re-sampling reference transects established during the earlier SCAT surveys using identical methods, by the same contractor (Dr. James Gibeaut), but is not included in this final report because the contractor has not submitted their report for this element. Presumably the contractor will submit a separate report for this element independently. [element will be added to the final report if completed in time]

Between 1989 and 1991, subtidal sediments were evaluated for oil contamination in transects from 0 m to 100 m seawater depths. In 2001, several "worst case" transects were re-sampled using identical sampling and hydrocarbon analysis methods, so the results of these efforts are directly comparable.

Chemical fingerprinting

The last objective, chemical fingerprinting, provided clear identification of the origins of the typical oil samples collected. This relied on analysis by gas chromatography/mass spectrometry of the aliphatic hydrocarbons and the polycyclic aromatic hydrocarbons (PAH), which definitively resolves EVO from other potential sources of oil in PWS. The most prominent other source of oil was derived from the Monterey Formation in California, as asphalt released during the 1964 Great Alaska Earthquake. When combined with the random sampling results, the fingerprinting results lead to estimates of the relative amounts of this "earthquake oil" in comparison with remaining EVO.

Results

The body of this report consists of three chapters, which present the results of the simple random sampling, the adaptive sampling, and the subtidal sediment sampling. The first chapter is being published in the peer-reviewed journal *Environmental Science and Technology*, and the other two chapters may be revised for publication elsewhere.

Sites with Oil

Oil was found in over half of the beach segments, using a quarter meter quadrat randomly placed in the upper half of the intertidal zone. Of the 91 beach segments sampled, oil was found in 53 segments. Eleven beaches had only surface oil, 14 had only subsurface oil, and 28 had both. The oil detection frequency ranged from 0% on 38 segments to over 30% on 6 others. The most heavily oiled segments were within sheltered embayments that received the brunt of the initial oil landfall. Other geomorphologic associations with oil detection frequency were less evident, although persistent oil was found on beaches that had boulder/cobble surface armoring, nearly level slopes of the middle intertidal, or a thick sediment veneer over a bedrock platform as reported previously.

Oil was also present on 25 of the 38 beaches where oil was not found in our randomly placed quadrats. This oil was found during surveys we performed using the SCAT methodology on each of the beaches visited to compare results from the two approaches directly. This oil was present as small (< 10 m^2) surface asphalt pavements or as tarballs in the upper intertidal discovered by our SCAT surveys (on 24 beaches), or as a large patch of subsurface oil discovered by our opportunistic sampling (1 beach). We found no evidence of oil at 13 beaches. Hence, we found some form of oil at 86% of the beaches we visited.

Estimate of Beach area contaminated with oil

Subsurface oil accounted for more contaminated beach area than did surface oil contamination. The results of the random sampling effort lead to an estimated 4.1 hectares (ha; 1 ha = 2.47 acres) of beach covered by surface oil and 7.8 ha of beach containing subsurface oil as of the summer of 2001. The 95% confidence intervals (CI) for these estimates is 2.1 - 7.7 ha for the surface oil, and 4.1 to 12.7 for the subsurface oil. The total beach oil that remained contaminated is almost the sum of these areas (11.3 ha), because surface and subsurface oil infrequently coincided. The surface oil was almost evenly distributed across the upper half of the intertidal, but the subsurface oil was detected more frequently toward the lower boundary of our sampling in the mid-intertidal. Subsurface oil was also found in the lower intertidal at 18 beaches, based on occasional pits that were excavated to substantiate the prevalence of lower intertidal oil. The oil observed below the grid was significant, but does not enter into the calculations of beach area contaminated or oil volume estimates.

The finding of large areas of subsurface oil in the lower part of our sampling grid, and below the grid in the lower intertidal were very surprising, in light of previous reports stating that most of the stranded oil persisted in the upper intertidal. A close reading of these earlier reports showed that the notion that the EVO persisted only in the upper intertidal was based on a conjecture that it would be most persistent there because it is drier, and hence the oil would be able to adhere better. This conjecture then guided subsequent sampling, producing results that appeared to validate it. But the conjecture was never rigorously evaluated until this project, and our results indicate it is substantially incorrect. The notion that the persistent oil was limited to the upper intertidal led to the conclusion that the oil probably had few effects on intertidal biota, which mainly inhabit the lower intertidal. This conclusion now requires reconsideration in light of the results of this project.

Our sampling underestimated the area of oil-contaminated beach in PWS because it excluded (i) tidal elevations lower than +1.8 m, (ii) beaches described as lightly or moderately oiled in 1989 but not thereafter, (iii) pit depths deeper than 0.5 m, and (iv) oil not evident visually or by odor. Of these, failure to sample the lower intertidal probably caused the greatest bias. The increasing frequency of subsurface oil from the upper (+4.8 m) to the mid-tidal (+1.8 m) elevation, along with the results of our opportunistic sampling of the lower intertidal suggests that subsurface oil may be encountered within the lower intertidal nearly as often as in the upper intertidal. The lower limit of the initial oil impact was about +0.5 m, based on the lowest low tide during the period of initial landfall after the spill, so the proportion of the initially oiled tidal elevation (+0.5 m to +4.8 m) below our sampling grid is about 30%, suggesting an underestimate of oiled beach of similar magnitude. In contrast, the small estimated contribution from the less heavily-oiled beaches we sampled that oiled beach areas in the unsampled portions of PWS are probably negligible, as are contributions from oil present at depths more than 0.5 m below the beach surface based on infrequent detection of oil in pits we occasionally dug to depths of 1 m. Considering all these sources of bias we believe it very unlikely that more than 25 ha. of the intertidal remained contaminated with surface or subsurface oil by summer 2001.

Estimate of the mass of oil remaining

We estimated the mass of subsurface oil by classifying oil into visual categories, and then measuring the mass of oil associated with typical occurrences of each category. Combination of these results with the oiled beach area associated with each visual oiling category permits estimation of the total mass of oil remaining. This estimate is 55,600 kg (1 kg = 2.2 lb), with a 95% CI ranging from 26,100 kg to 94,400 kg. Consideration of the extensive asphalt pavements of surface oil encountered, which were up to 3 cm thick, and the unsampled portions of PWS leads us to conclude that the total amount of oil mass is probably not more than twice our estimate of subsurface oil (i.e. < 190,000 kg).

Adaptive sampling estimates

Results from the adaptive sampling estimators are slightly (but not significantly) lower, because different estimation formulas and data are used for the adaptive sampling estimates. The adaptive sampling estimates are also slightly more precise, but the increased precision was not justified by the effort expended. In retrospect, our sampling would have been more precise had we sampled more beaches instead of sampling adaptively.

Although the volume of oil has declined considerably, our study suggests the area of oiled beach has probably changed little since 1992. Comparison of subsurface oiling intensities found during our study suggests that dispersion after 1992 may not have reduced the oiling intensity enough to reduce the area of visibly oiled beach. Our 2001 estimates of surface and subsurface oil are higher than the highest results reported for 1992 or 1993 by factors of two or more. The SCAT methods used for the earlier studies underestimated the extent of oiled beach somewhat because some oiled beaches were not sampled, and because sampling effort was directed mainly toward the upper intertidal where oil was incorrectly thought to be most persistent. However, to conclude that the area of oiled beach significantly diminished from 1993 to 2001, the earlier survey results would require adjustment upwards by a factor of three or more before exceeding the upper bounds of our confidence intervals, and we believe an adjustment of this magnitude is unrealistic in view of the large sampling effort of the previous surveys.

Although the oil remaining is only about 0.14% - 0.28% of the volume originally beached, the decline was most rapid during the first few years. About 2% of the original spill volume of *Exxon Valdez* oil was estimated to remain on PWS beaches by fall, 1992, implying losses of about 58% per year during the first 3.5 years. Comparison with our 2001 estimate of 55,600 kg, possibly an underestimate by a factor of two, implies an annual loss rate of 20% - 26%, substantially slower than anticipated.

Chemical fingerprinting

All of the subsurface oil fingerprinted back to EVO. Nearly all of the surface oil was EVO (-90%). The few visible oil occurrences that did not match EVO were Monterey Formation asphalts, and probably account for < 10% of the surface oil. This confirms that EVO remains by far the largest reservoir of toxic PAH on beaches most impacted by the *Exxon Valdez* oil spill in PWS.

Subtidal oil

We sampled inter- and subtidal sediments from one reference and four oiled locations in PWS to evaluate the persistence of remnant, subtidal oil from the 1989 EVOS. At each location, sediment samples were collected from seven depths along transects beginning at 0 m and ending at 100 m, at stations that had been repeatedly sampled during surveys from 1989 to 1991. The collected sediments were analyzed for a suite of aliphatic and polycyclic aromatic hydrocarbons (PAH) using methods directly comparable with those used during the 1989 - 1991 surveys.

Oil concentrations found in subtidal surface sediments were very low, lower than the concentrations of remnant *Exxon Valdez* oil found in 1989-91 or in the subsurface pits found in 2001 in the intertidal zone. Subtidal oil concentrations from the oiled and reference locations were similar, and showed PAH concentrations that generally increased with depth from total PAH concentrations near detection limits of a few ng/g at 0 m to 776 ng/g at 100 m. Most of these PAH were from the regional background of coal and organic-rich shales that erode from the terrestrial Kulthieth Formation east of the Copper River and transported into PWS by the Alaska Coastal Current. Small contributions from anthropogenic sources were likely present at three stations, and trace contributions from *Exxon Valdez* oil may be present at two, but the evidence for this is weak. Because the four locations we sampled in western PWS during 2001 were among the most heavily contaminated by the Exxon Valdez oil spill, we believe contributions from remnant *Exxon Valdez* oil to the subtidal sediments are now negligible in comparison with contributions from the regional background.

Conclusion

We found more oil than expected 12 years after the Exxon Valdez oil spill. Oil was found on 77 of 91 beaches randomly selected according to their oiling history. Surface oiling was recorded for randomly placed quadrats, which were then excavated and examined for subsurface oil. The cumulative area of beach contaminated by surface or subsurface oil was estimated at 11.3 ha. Surface oil varied little with tide height, but subsurface oil was more prevalent at the lower tide heights. The mass of remaining subsurface oil is conservatively estimated at 55,600 kg. Analysis of terpanes indicated all of the subsurface oil and over 90% of the surface oil was from the *Exxon Valdez.*

The unexpected persistence of subsurface *Exxon Valdez* oil, often only moderately weathered and extending into the more biologically productive middle and lower intertidal, confirms the potential for long-term exposure risk and biological effects to biota dependent on these beaches.

An Estimate of Oil persisting on Beaches of Prince William Sound, 12 years after the Exxon Valdez oil spill

Jeffrey W. Short*, Mandy R. Lindeberg, Patricia M. Harris, Jacek M. Maselko, Jerome J. Pella, and Stanley D. Rice

Auke Bay Laboratory, Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA, 11305 Glacier Highway, Juneau, Alaska 99801-8626

*Corresponding author telephone: 907-789-6065; fax: 907-789-6094; e-mail: Jeff.Short@noaa.gov

Abstract

We estimated the amount of oil remaining in Prince William Sound, Alaska, 12 years after the 1989 Exxon Valdez spill to assess its importance as a long term reservoir of toxic hydrocarbons. We found oil on 77 of 91 beaches randomly selected according to their oiling history. Surface oiling was recorded for randomly placed quadrats, which were then excavated and examined for subsurface oil. The cumulative area of beach contaminated by surface or subsurface oil was estimated at 11.3 ha. Surface oil varied little with tide height, but subsurface oil was more prevalent at the lower tide heights sampled. The mass of remaining subsurface oil is conservatively estimated at 55,600 kg. Analysis of terpanes indicated over 90% of the surface oil and all of the subsurface oil was from the *Exxon Valdez*, and that Monterey Formation oil deposited after the 1964 Alaska earthquake account for all the remaining surface oil. These results indicate that oil from the *Exxon Valdez* remains by far the largest reservoir of biologically available polycyclic aromatic hydrocarbons on beaches impacted by the spill, and that biota dependent on these beaches risk continued exposure.

Introduction

Shorelines polluted by oil spills require practical methods for monitoring their impacts. The extent of oiling must be monitored for effectively allocating beach protection and clean-up resources, for determining persistence, and for assessing damages. These determinations must be done quickly, reliably and inexpensively, often in an atmosphere of crisis. The method usually used relies on shoreline cleanup assessment teams (SCAT), who comprehensively inspect beaches and record results following standardized procedures, hereafter termed the SCAT method (*1 – 4*). Excavations to monitor subsurface oil, especially at longer time intervals following an incident, when subsurface oil may not be obvious, may augment these visual assessments. The SCAT method produces a point estimate of the oiled area and a qualitative indication of oiling intensity; these tend to be underestimates because some subsurface oil may escape detection.

The SCAT method, used extensively for four years following the 1989 *Exxon Valdez* oil spill in Prince William Sound (PWS), provided the basis for defining the initial extent of oiled beaches, and for monitoring oil persistence by both government- and industry-supported teams (*5*, *6*). According to a 1992 SCAT survey, the extent of oiled shoreline had decreased from 783 km to 10 km (*6*), with most of the oil, including subsurface oil, in the upper intertidal or supratidal zones. Based on excavations of thousands of pits from 1990 to 1992, the cumulative area of beach contaminated by subsurface oil had diminished to an estimated 12,000 m^2 (*6*). These trends suggested that oil remaining after 1992 would soon disperse to negligible amounts (*5 – 9*).

At some locations, *Exxon Valdez* oil seemed more persistent than expected based on the 1990-1992 SCAT surveys. A survey conducted in 1993 suggested that remaining oil might persist because it was no longer disturbed by cleanup efforts (*5*). Where oil was protected from dispersion by cobble-boulder surface armor (*10*, *11*) or by mussel beds (*12*), subsurface oil was often encountered through the late 1990s. A beach cleaning effort at Sleepy Bay in 1997 encountered substantial deposits of subsurface oil (*13*), which raised questions regarding oil persistence elsewhere. Public concerns regarding the extent of remaining oil led the Exxon Valdez Trustee Council, a consortium of federal and State of Alaska agencies, to support the study we report here, conducted during summer 2001.

We used methods based primarily on random sampling, instead of the comprehensive assessment approach of the SCAT method, because we wanted to provide a quantitative, probability-based estimate of the amount of oil remaining, (i.e., an interval estimate as well as a point estimate). We were not confident that the surface oil remaining in 2001 would be a reliable indicator of subsurface oil, which also indicated randomization of sampling effort. A SCAT survey of surface oil was conducted on each of our randomly selected beaches so that results from our randomly placed quadrats could be compared. Our objectives were to provide estimates and confidence intervals for the total area of beach that remained contaminated in PWS and for the mass of oil remaining.

We employed two compatible random sampling designs for estimating oiled areas on individual beach segments: stratified random sampling (SRS) and stratified random adaptive sampling (SRAS). These involved a large-scale sampling effort with excavation of nearly 9,000 pits. The

SRAS design gives more precise estimates when sampling something that is rare and highly aggregated, and was included in case the detection frequency of oil was very low. We report here results from the SRS design only, because the detection frequency was not rare, both designs produced estimates that did not differ materially, and the increased precision of the SRAS design was negligible. The SRAS results are available from the authors on request.

Methods

Our study area included all the shorelines in the western PWS impacted by the *Exxon Valdez* spill (Figure 1). We sampled three mutually exclusive categories of beaches defined by the persistence of visually evident oil during SCAT surveys conducted from 1989 through 1993 (*5*, *14*). These categories included discrete segments of beaches that were described as heavily (Category I) or moderately (Category II) oiled at any time during the period 1990 to 1993, and beaches that were described as heavily oiled during 1989 but had only light to no oil impact during subsequent years (Category III). Category III beaches were not expected to have significant amounts of oil, but surveys after 1989 were scant for this category. The total length of all the segments is 116.6 km, comprising lengths of 24.4 km, 49.1 km, and 43.1 km in categories I, II, and III, respectively.

Sampling Design

Our field crew was able to sample one 100-m length of beach during one low tide interval, so our basic sampling unit for beaches is segments of 100 m or less. Shorter segments resulted from shorter beach lengths described during the earlier surveys, or from the remainders of long beach segments after partitioning into 100 m segments. This resulted in two classes of beach segments (i.e. 100 m and < 100 m) and six sampling strata, the product of the three sampling categories and two segment classes. Beach segments in each category were selected by simple random sampling without replacement if 100 m in length, or by probability proportional to length (PPL) if less than 100 m. The PPL sampling requires sampling with replacement. A total of 7.9 km of shoreline, comprising 91 distinct segments (Figure 1), was selected, and sampled with 5.3 km from Category I where most remaining oil was anticipated. Cumulative lengths of Category II and III were 1.9 km and 0.75 km, respectively. The proportion of available beach that was selected within the < 100 m and 100 m classes were, respectively, 26.6% and 20.1% in Category I, 7.6% and 2.8% in Category II, and 2.2% and 1.6% in Category III.

The selected beach segments were located by global positioning system (GPS) identification of the selected segment coordinates after reconciliation with known landmarks and data recorded on maps of previous surveys (*14*). Tidal elevation was determined by an infrared laser-level beam, the height of which was determined by survey reference to observed tide height and predicted tide heights for Cordova, Alaska, corrected for the nearest tidal monitoring station. We estimate errors associated with planar coordinates as about ± 2 m, based mainly on GPS uncertainty, and with vertical elevations as about ± 0.25 m, because we did not correct observed tide heights for atmospheric pressure.

We established a sampling grid between +1.8 m and +4.8 m above tidal datum on each beach segment (Figure 2). Each 100-m beach section length was divided into eight columns, each 12.5 m wide. Shorter beach segments were divided into fewer columns. A meter-tape laid alongshore on the beach and other meter-tapes laid perpendicular along the column boundaries delineated the grid boundaries. Beaches were searched visually for evidence of surface oil within the grid following the SCAT protocols. Leaders of previous SCAT surveys in PWS trained two of us in the proper identification of the SCAT oiling classifications in 1993 (*5*). The areal extent of surface oil was estimated as proportions of beach area covered by reference to the meter tapes, and each surface oil patch was classified according to type: asphalt pavements/mousse (AP/MS, classes combined here), surface oil residue (SOR), tar balls (TB), coat (CT), or oil film (OF) (*5*).

The grids were sampled by SRS concurrently with the SCAT surveys. The sampling grid columns were partitioned into rectangular blocks by six 0.5-m vertical tidal elevation intervals, resulting in 48 blocks on a 100-m beach (Figure 2). Two 0.25-m^2 quadrats were randomly placed within each block, based on randomly selected distances along the meter tapes. Each quadrat was evaluated for the presence of surface and subsurface oil. Oil visually evident within the uppermost 5 cm of a beach surface was considered surface oil. The presence of subsurface oil was evaluated by digging a test pit within each quadrat to a depth of 0.5 m or until boulders or bedrock intervened, and examining the pit for evidence of oil by sight and smell. Subsurface classifications included no oil (NO), oil film (OF), light, medium, or heavy oil residue (LOR, MOR, and HOR respectively), as defined in ref. (*5*). A total of 7,484 random quadrats were drawn, of which 6,775 were excavated, with the remainder located on surface boulders, bedrock, or cliffs. We occasionally dug pits deeper than 0.5 m, or at intertidal elevations below the lower grid boundary of +1.8 m to evaluate whether oil occurred outside of our sampling design boundaries.

Estimation of Oiled Beach Areas

The estimated total area of beach covered by oil discovered during our SCAT surveys is simply the product of the measured oiled area multiplied by the ratio of available and selected beach lengths for each of our six beach-sampling strata. Results within categories are summed for presentation here (cf. Table 1).

The oiled area A_j for the jth sampled beach segment, based on the randomly placed quadrats, is estimated as

$$\hat{A}_j = \sum_{i=1}^{K} N_i \bar{y}_i, \qquad (1)$$

where K is the number of blocks in the entire segment, N_i is the ratio of the area of block i and the area of the quadrat (0.25 m^2), and $\bar{y}_i$ is the average number of oiled quadrats found in block i (either 0, 0.5, or 1).

The two classes of beach segments (i.e., 100-m and <100-m lengths) had different estimators of oiled area. The total oiled area $T_{s,100m}$ for each of the three 100-m categories is estimated as

$$\hat{T}_{s,100m} = \frac{N_{100m}}{n_{100m}} \sum_{j=1}^{n_{100m}} \hat{A}_j, \qquad s = 1,2,3 \qquad (2)$$

where N_{100m} is the total number of beach segments in each category s within the 100-m oiling category, n_{100m} is the sampled number of beach segments, and $\hat{A}_j$ is the estimated oiled area in the jth sampled beach segment.

The total oiled area $T_{s,<100m}$ for each of the three <100-m categories is estimated as

$$\hat{T}_{s,<100m} = \frac{L_{<100m}}{n_{<100m}} \sum_{j=1}^{n_{<100m}} \frac{\hat{A}_j}{L_j}, \qquad s = 1,2,3 \qquad (3)$$

where $L_{<100m}$ is the total length of beach segments in each category s within the < 100-m oiling category, $n_{<100m}$ is the sampled number of beach segments, $L_{j<100m}$ is the length of the jth beach segment, and $\hat{A}_j$ is the estimated oiled area in the jth sampled beach segment.

The total oiled area in Prince William Sound, T_{PWS} is the sum of the six oiled area estimates:

$$\hat{T}_{PWS} = \sum_{s=1}^{3} (\hat{T}_{s,100m} + \hat{T}_{s,<100m}) \qquad (4)$$

Surface and subsurface oiled areas were estimated separately in this fashion.

Gravimetric Analysis of Oil

Oil was extracted from typical samples of each subsurface oil classification and weighed to establish a basis for estimating the mass of oil remaining in PWS. A total of 97 test pits, 15 HOR, 26 MOR, 45 LOR and 11 NO, each $(0.5\ m)^3$ were selected and sampled on a schedule of 0 – 5 pits per day throughout the field season in order to obtain broad coverage of PWS beaches for the three categories. Oil was scraped from particles larger than 6 cm and added to the remaining contents of the pit, which were weighed and thoroughly mixed, after which a 2 kg subsample was extracted twice, each with 1 L of dichloromethane for 2 h. The dichloromethane of the combined aliquots was distilled over steam and the residual material weighed. Results are expressed as kg oil m^{-2} beach surface area, where the oil mass is determined as the product of the oil mass per kg in the 2-kg subsample and the mass of the homogenized test pit contents.

Estimation of Subsurface Oil Mass

To estimate the mass of oil in each beach segment, we calculated the proportion of oiled area belonging to each of the three visual oiling categories, HOR, MOR, and LOR, and then multiplied the proportion by the average weight of oil in that category. The mass of the NO category was assumed to be zero. The total oil mass estimate for an individual beach segment is the sum of the oil masses for the three visual oiling categories.

The total mass of oil in each of the three categories of the 100-m beach segment class is denoted as $W_{s,100m}$, and estimated as

$$\hat{W}_{s,100m} = \frac{N_{100m}}{n_{100m}} \sum_{j=1}^{n_{100m}} \sum_{m=1}^{3} \hat{A}_j \hat{p}_{jm} \overline{w}_m, \qquad s = 1,2,3 \qquad (5)$$

where $\hat{A}_j$ is the estimated total oiled area of the jth beach segment in category s, $\hat{p}_{jm}$ is the proportion of oiled quadrats belonging to visual oiling category m within beach segment j, and $\overline{w}_m$ is the average mass of oil in visual oiling category m estimated from gravimetric samples.

The total mass of oil in each of the three categories of the <100-m beach segment category is denoted as $W_{s,<100m}$, and estimated as

$$\hat{W}_{s,<100m} = \frac{L_{<100m}}{n_{<100m}} \sum_{j=1}^{n_{<100m}} \sum_{m=1}^{3} \frac{\hat{A}_j}{L_j} \hat{p}_{jm} \overline{w}_m , \quad s = 1,2,3 \qquad (6)$$

where $\hat{A}_j$ is the estimated total oiled area of the jth beach segment in category s, $\hat{p}_{jm}$ is the proportion of oiled quadrats belonging to visual oiling category m within beach segment j, and $\overline{w}_m$ is the average mass of oil in visual oiling category m estimated from gravimetric samples.

The total mass of oil in all of Prince William Sound, W_{PWS} is estimated as

$$\hat{W}_{PWS} = \sum_{s=1}^{3} (\hat{W}_{s,100m} + \hat{W}_{s,<100m}) . \qquad (7)$$

We used bootstrap-generated distributions from 1,000 resamplings of masses or areas to estimate standard errors and 95% confidence intervals (95% CI) per stratum. Resampling with replacement was used for the strata with <100-m segments and resampling without replacement for the 100-m segment strata. Precision of total oiled areas and masses across strata was determined similarly from 1,000 totals of estimates per stratum.

Qualitative Analysis of Oil

Typical samples of surface (n = 27) and subsurface (n = 38) oil deposits along with 7 samples of unoiled (control) sediments were analyzed by gas chromatography-mass spectrometry (GCMS) to evaluate oil sources and weathering. The surface oil samples included 15 AP/MS samples, 3 SOR samples, and 9 tarball samples. The subsurface oil samples included 6 HOR samples, 13 MOR samples, 18 LOR samples, and 1 OF sample. Bulk tarballs and asphalt pavement samples (1 mg – 100 mg) were dissolved in dichloromethane. Sediment samples (0.1 g – 7 g) were extracted with dichloromethane at 100 °C and 2000 psi for 10 min in a Dionex ASE 200 accelerated solvent extractor. The dichloromethane solutions were exchanged with hexane over steam, and separated into aliphatic and aromatic fractions by column chromatography (10 g 2%-deactivated alumina over 20 g 5%-deactivated silica gel). Aliphatics eluting with 50 mL pentane were analyzed by GCMS at m/z 191 for terpanes (*15*) following concentration to ~ 1 mL hexane over steam. Polycyclic aromatic hydrocarbons (PAH) eluting with 50 mL 1:1 (v/v) pentane:dichloromethane were exchanged into ~ 1 mL hexane over steam, and further purified by size-exclusion high-performance liquid chromatography. Purified PAH were measured by GCMS operated in the selected ion monitoring mode (*16*). Analyzed PAH include parent and alkyl-substituted naphthalenes, fluorenes, dibenzothiophenes, phenanthrenes, and chrysenes.

Oil source identification is by comparison with reference chromatograms of terpanes for *Exxon Valdez* oil and for asphalt originating from the Monterey Formation in California, which contaminated PWS following release from storage tanks during the 1964 Alaska earthquake (*15*). *Exxon Valdez* oil was identified on the basis of the joint absence of 17α(H),18α(H),21β(H)-28,30-bisnorhopane and of 18(α+β)(H)-oleanane, and ratios near 2 of Triplet (C_{26}-tricyclic terpane (S?) + C_{26}-tricyclic terpane (R?))/C_{24}-tetracyclic terpane) (*17*). Oil derived from the Monterey Formation contained bisnorhopane and oleanane, and had ratios near 4 of Triplet (*17*). The weathering state of *Exxon Valdez* oil is indicated by an index, *w*, of PAH weathering losses (*18*). Slightly weathered oil is indicated by $0 < w < 2$, moderate weathering by $2 < w < 6$, and heavy weathering by $w > 6$.

Results

Proportion of Beaches Still Oiled

The distribution of detected oil among sampled beaches was highly variable. We found oil in randomly placed quadrats on 53 of the 91 selected beach segments. Eleven had only surface oil, 14 had only subsurface oil, and 28 had both. The oil detection frequency ranged from 0% on 38 segments to over 30% on 6 others. The most heavily oiled segments are within sheltered embayments that received the brunt of the initial oil landfall. Other geomorphologic associations with oil detection frequency were less evident, although persistent oil was found on beaches that had boulder/cobble surface armoring, nearly level slopes of the middle intertidal, or a thick sediment veneer over a bedrock platform as reported previously (*11*).

Oil was also present on 25 of the 38 beaches where oil was not found in our randomly placed quadrats. This oil was present as small (< 10 m^2) surface asphalt pavements or as tarballs in the upper intertidal discovered by our SCAT surveys (on 24 beaches), or as a large patch of subsurface oil discovered by our opportunistic sampling (1 beach). We found no evidence of oil at 13 beaches, 7 of which were from category III. Hence, we found some form of oil at 86% of the beaches we visited, and on 93% of the combined beaches of Categories I and II.

Surface Oil Estimates

The beach area covered by surface oil found by the SCAT surveys was just under 1 ha (Table 1). Oil was found in all three beach categories, with about 70% on the Category I beaches. The 100 m-segments accounted for 71% of the oil detected. Overall, this oil was mainly present as AP/MS (55%) and SOR (39%), followed by TB (2.9%), CT (2.0%), and OF (0.66%).

The SRS estimate of surface oiled beach area is 4.13 ha (95% CI: 2.07 ha - 7.05 ha; Table 2), which is significantly larger than the SCAT-based estimate. This area is nearly evenly divided between Category I and Category II beaches, and more oil was found in the 100-m segments than the <100-m segments. Surface oil was not detected within the randomly placed quadrats on any of the Category III beaches sampled, in contrast with the SCAT survey results.

The 100-m segments of Category II accounts for 91% of the variance of the total area estimate, a consequence of the lower proportions of area sampled compared with the other selection classes of Categories I and II. The SCAT-based estimates fall just under the lower bounds of the respective 95% CIs for the 100-m and the <100-m segments of Categories I and II (compare Tables 1 and 2).

The close scrutiny of the randomly placed SRS quadrats revealed surface oil that was not found during the SCAT surveys on 13 beaches, compared with the 24 beaches where surface oil was found by the SCAT surveys but not within the SRS quadrats.

The surface oil was distributed throughout the upper half of the intertidal. Surface oil was most often detected within the +3.3 m – +3.8 m tidal elevation stratum, but was also frequently detected near the upper and lower boundaries of our sampling grid (Figure 3-A).

Subsurface Oil Estimates

The SRS estimate of beach area contaminated by subsurface oil was 7.80 ha (95% CI: 4.06 ha – 12.7 ha; Table 3). As with the SRS-estimates of surface oil, the subsurface oiled area is nearly evenly divided between Category I and Category II beaches, with less than 10% in Category III beaches, and most of the variance (78%) is from the 100-m segments of Category II. Subsurface oil was most often encountered as LOR (62%), followed by MOR (21%), OF (11%), and HOR (6%).

Subsurface oil was more often detected in the lower tidal elevations of our sampling grid (Figure 3-B). The trend of increasing frequency of subsurface oil at lower tidal elevations suggests that oil may often be found in the lower intertidal, corroborated by our discovery of subsurface oil at tidal elevations below our sampling grid, often extending to 0 m tidal elevation, on 18 beaches during our opportunistic sampling.

Our SRS estimate of beach surface area contaminated by either surface or subsurface oil is 11.3 ha (95% CI: 6.78 ha – 17.2 ha). This estimate is not simply the sum of surface and subsurface area estimates because these areas intersect. Overall, the intersection is estimated as 15.7% of the surface oiled area, and is lower for Category I beaches (10.8%) than for Category II beaches (20.2%).

Oil Mass

The mass of oil per unit surface area in the gravimetric samples was variable, but the means roughly double between categories in ascending the oil classifications (Figure 4). The mean mass of oil per unit area of beach ranged from 0.59 kg m^{-2} for LOR to 2.1 kg m^{-2} for HOR. Incorporation of these results into eqs 5, 6, and 7 led to an estimate of 55,600 kg (95% CI: 26,100 kg – 94,400 kg; Table 4). Oil mass variability estimates are consistently somewhat greater than those for corresponding subsurface oiled areas (compare Tables 2, 3, and 4), but follow similar trends.

Oil Sources and Weathering

All of the surface oil samples analyzed were either *Exxon Valdez* oil or Monterey Formation asphalt. All nine of the tarball samples were derived from the Monterey Formation, with oleanane and bisnorhopane clearly evident, and Triplet ratios from 2.7 to 4.5. All but 2 of the 14 AP/MS samples were *Exxon Valdez* oil, with oleanane and bisnorhopane absent and Triplet ratios from 0.83 to 2.20. The two AP/MS samples from the Monterey Formation were small (< 0.25 m^2) tar mats collected from the upper intertidal. Two of the *Exxon Valdez* AP/MS samples were slightly weathered, three were moderately weathered and the rest were heavily weathered, with $w > 10$ for six of the samples. All three SOR samples were moderately weathered *Exxon Valdez* oil.

All of the 38 subsurface oil samples had Triplet ratios from 1.04 to 2.18 and lacked bisnorhopane and oleanane, consistent with *Exxon Valdez* oil. The oil in these samples was usually less weathered than the surface oil samples, with the median $w = 3.3$ (range 0.94 – 12.1).
The sum of the PAH concentrations in the seven control samples and the OF sample range from 19 ng/g to 100 ng/g, compared with concentrations ranging from 355 ng/g to 14.7 mg/g (median

68.6 μg/g) in other sediment and oil samples.

Discussion

Oil Sources

Surface oil in the sampled region came from the *Exxon Valdez* spill, or from storage tanks of asphalt and other Monterey Formation petroleum products that ruptured during the 1964 earthquake (*15*). Oil from the Monterey Formation was usually found above +3-m tide height and typically occurs as flattened tarballs firmly adhered to cobbles and boulders, and infrequently as small (<0.25 m^2) tar mats (*15*, this study). This distribution suggests that tarballs were stranded by high tides and softened sufficiently during periods of high insolation to dry and adhere to rocks before the next high tide. This stranding mechanism would limit the size of tarballs because large tarballs, less likely to absorb enough heat to adhere, would be more susceptible to further waterborne transport and ultimately would exit into the Gulf of Alaska. Oil from the *Exxon Valdez* was much less viscous than commercial asphalt, and initially came ashore under high-energy wave conditions throughout the range of tidal excursion, leading to very extensive oil patches throughout the intertidal area. This led to greater variety in the occurrences of persistent *Exxon Valdez* oil, including subsurface oil, extensive asphalt pavements, sheening surface oil residues, and possibly tarballs that may be difficult to distinguish from Monterey Formation tarballs in the field (*15*).

We were unable to chemically characterize each of the hundreds of oil occurrences we encountered during this project, so instead we collected type samples based on visual appearance. All of the tarballs and two small asphalt pavement samples were from the Monterey Formation source, but the larger asphalt pavement samples (>0.25 m^2) we encountered were from *Exxon Valdez* oil. Because the tarballs as a whole account for < 3% of the surface oil encountered in our SCAT survey, we conclude that petroleum derived from the Monterey Formation probably accounts for less than 10% of the surface oil we discovered. Because all of the subsurface oil is *Exxon Valdez* oil, Monterey Formation oil remains a minor source of hydrocarbons on the beaches most heavily impacted by the *Exxon Valdez* spill. Other recent anthropogenic sources have been suggested as important based on anecdotal information (*19*), but we did not encounter any evidence for these sources on any of our 91 beaches.

Accuracy, Precision and Bias of Estimates

Our sampling underestimated the area of oil-contaminated beach in PWS because it excluded (i) tidal elevations lower than +1.8 m, (ii) beaches described as lightly or moderately oiled in 1989 but not thereafter, (iii) pit depths deeper than 0.5 m, and (iv) oil not evident visually or by odor. Of these, failure to sample the lower intertidal probably caused the greatest bias. The increasing frequency of subsurface oil from the upper (+4.8 m) to the mid-tidal (+1.8 m) elevation grid limits (Figure 3-A), along with the results of our opportunistic sampling of the lower intertidal suggests that subsurface oil may be encountered within the lower intertidal nearly as often as in the upper intertidal. The lower limit of the initial oil impact was about +0.5 m, based on the lowest low tide during the period of initial landfall after the spill, so the proportion of the initially oiled tidal elevation (+0.5 m to +4.8 m) below our sampling grid is about 30%, suggesting an underestimate of oiled beach of similar magnitude. In contrast, the small estimated contribution from category III beaches suggests that oiled beach areas in the unsampled portions of PWS are probably negligible, as are contributions from oil present at depths more than 0.5 m below the beach surface based on infrequent detection of oil in pits we occasionally dug to depths of 1 m. Considering all these sources of bias we believe it very unlikely that more than 25 ha. of the intertidal remained contaminated with surface or subsurface oil by summer 2001.

Our estimate of oil mass is directly affected by underestimation of the oiled area, and also by exclusion of surface oil. Assuming a mean surface oil thickness of 1 mm (and some asphalt pavements were 3 cm thick), an oiled area of 4 ha. implies 40,000 L, or about 36,000 kg of surface oil. The amount of oil that remained is therefore probably greater than our estimate of 55,600 kg, perhaps by as much as a factor of two (accounting for both the subsurface oil in the lower intertidal as well as the surface oil).

Our estimates would have been more precise if we had allocated more sampling effort to the Category II, 100-m beach segments at the expense of our adaptive sampling effort. Our estimates of oiled areas on individual beaches contributed almost negligible variance compared to variance associated with expanding these estimates to the whole of the respective sampling classes. The greatest uncertainty arose from the Category II, 100-m sampling class, where we sampled only 2.8% of the class and where substantial oil was occasionally encountered.

Comparison of 2001 SCAT and Random Sampling Methods

The consistently larger estimates of beach area contaminated by surface oil derived from the SRS compared with the SCAT method is primarily due to two factors. First, the definition of surface oil used for the SRS method includes oil within the upper 5 cm of beach sediments, and this was not always evident without disturbing the surface material. Second, surface oil was often obscured by weathering or epiflora from casual observation. The SRS method involved closer scrutiny of small (0.25 m^2) portions of beach, so less obvious surface oil was more likely to be observed.

The detection of some surface oil in the Category III, 100-m class by the SCAT method, but not by the SRS method, is a stochastic result of the random allocation of the SRS sampling quadrats. Oil detected during the SCAT surveys was not detected within the SRS quadrats on 10 other beaches as well, and the possibility of not finding oil within the SRS quadrats when oil patches are in fact present within the sampling grid contributes to the variance of the estimates of total oiled areas.

Oil Persistence

Although the volume of oil has declined considerably, our study suggests the area of oiled beach has probably changed little since 1992. Comparison of subsurface oiling intensities found during our study (mainly LOR) with intensities reported for the previous studies (mainly MOR and HOR; *5*, *6*) suggests that dispersion after 1992 may not have reduced the oiling intensity enough to reduce the area of visibly oiled beach. Our 2001 estimates of surface and subsurface oil are higher than the highest results reported for 1992 (*6*) or 1993 (*5*) by factors of two or more. The methods used for the earlier studies underestimated the extent of oiled beach somewhat because some oiled beaches were not sampled, and because sampling effort was directed mainly toward the upper intertidal where oil was incorrectly thought to be most persistent. However, to conclude that the area of oiled beach significantly diminished from 1993 to 2001, the earlier survey results would require adjustment upwards by a factor of three or more before exceeding the upper bounds of our confidence intervals, and we believe an adjustment of this magnitude is unrealistic in view of the large sampling effort of the previous surveys.

Although the oil remaining is only about 0.14% - 0.28% of the volume originally beached, the decline was most rapid during the first few years. About 2% of the original spill volume of *Exxon Valdez* oil was estimated to remain on PWS beaches by fall, 1992 (*20*), implying losses of about 58% per year during the first 3.5 years. Assuming a density of 0.96 for weathered oil (*21*), this volume is equivalent to about 806,000 kg. Comparison with our 2001 estimate of 55,600 kg, possibly an underestimate by a factor of two, implies an annual loss rate of 20% - 26%, substantially slower than anticipated (*9*).

The persistence of subsurface oil in the mid- and lower-intertidal area was unexpected. Previous work emphasized the prevalence of persistent oil in the upper intertidal (*5*, *6*, *10*, *11*, *22*), based on the conjecture that oil would be most persistent in the upper intertidal because it is drier there and so would be more adhesive (*3*). Subsequent sampling focused on the upper intertidal (*5*, *6*), producing results that appeared to confirm the conjecture. Although surface oil may very well have persisted mainly in the upper intertidal, our study indicates the same is not true for subsurface oil. In addition, the middle and especially the lower intertidal are intrinsically less accessible, and sampling there is more dangerous because the often extensive epiflora can be extremely slippery to traverse. We therefore suspect that subsurface oil in the middle and lower intertidal often escaped detection during previous surveys because of inadequate sampling effort.

The distributions of surface oil and of subsurface oil were clearly different by 2001, but the processes causing this divergence are not clear. The criteria used for the SCAT surveys of 1989 – 1993 did not differentiate among a wide range of oiling intensities that were lumped as "heavy", which might include oiled sediment lenses ranging from 1 cm to 1 m in thickness. Hence, the distribution of extremely oiled sediments was not documented, and the distribution of the persistently oiled sediments we found during 2001 may reflect the initial distribution of extremely oiled sediments.

The unexpected persistence of subsurface *Exxon Valdez* oil, often only moderately weathered and extending into the more biologically productive middle and lower intertidal, confirms the potential for long-term biological effects after 1992 on beaches most heavily impacted by the spill. While adequate for allocating oil spill response measures, our study indicates the usual SCAT method may not be adequate for monitoring the long term persistence of oil quantitatively, at least in cases such as the *Exxon Valdez* where a considerable proportion of beached oil may not be obvious after a few months.

Acknowledgements

The findings and conclusions presented by the authors are their own and do not necessarily reflect the views or position of the Trustee Council.

Literature Cited

1. Finkelstein, K.; Gundlach, E. R. *Environ. Sci. Technol.* **1981**, *15*, 545–549.

2. Owens, E. H. *Mar. Pollut. Bull.* **1987**, *18*, 110–118.

3. Owens, E. H., In *Proceedings, 14th Arctic and Marine Oil Spill Program (AMOP) Technical Seminar*, Environment Canada, Ottawa, ON, 1991, pp. 579–606.

4. Owens, E. H., In *Proceedings, 22nd Arctic and Marine Oil Spill Program (AMOP) Technical Seminar*, Environment Canada, Ottawa, ON, 1999, pp 337–360.

5. Gibeaut, J. C.; Piper, E. 1993 Shoreline Oiling Assessment of the Exxon Valdez Oil Spill. EVOS Restoration Project Final Report 93038. Exxon Valdez Trustee Council, Anchorage, AK, 1998.

6. Neff, J. M.; Owens, E. H.; Stoker, S. W. In *Exxon Valdez Oil Spill: Fate and Effects in Alaskan Waters*; Wells, P.G., Butler, J. N., Hughes, J. S., Eds.; American Society for Testing and Materials Pub. 1219: Philadelphia, Pennsylvania, 1995; pp 312–346.

7. Gilfillan, E. S.; Page, D. S.; Neff, J. M.; Parker, K. R.; Boehm, P. D. In *Proceedings, 2001 International Oil Spill Conference*, American Petroleum Institute, Washington, D.C., 2001, pp 559–567.

8. Koons, C. B.; Jahns, H. O. *Mar. Technol. Soc. J.* **1992**, *26*, 61–69.

9. Boehm P. D.; Page, D. S.; Gilfillan, E. S.; Stubblefield, W. A.; Harner, E. J. In *Exxon Valdez Oil Spill: Fate and Effects in Alaskan Waters*; Wells, P. G., Butler, J. N., Hughes, J. S., Eds.; American Society for Testing and Materials Pub. 1219: Philadelphia, Pennsylvania, 1995; pp. 347–397.

10. Hayes, M. O.; Michel, J. National Oceanic and Atmospheric Administration Technical Memorandum NOS ORCA 126, Washington, D.C., 1998, 115 p.

11. Hayes, M. O.; Michel, J. *Mar. Pollut. Bull.* **1999**, *38*, 92–101.

12. Carls, M. G.; Babcock, M. M.; Harris, P. M.; Irvine, G. V.; Cusick, J. A.; Rice, S. D. *Mar. Environ. Res.* **2001** *51*, 167–190.

13. Brodersen, C., Short, J., Holland, L., Carls, M., Pella, J., Larsen, M., Rice, S. In *Proceedings, 22nd Arctic & Marine Oilspill Program (AMOP) Technical Seminar*, Environment Canada, Ottawa, ON, 1999, pp 325–336.

14. Anonymous. EVOS Geographic Information System (GIS) Database and Data Dictionary. EVOS Research and Restoration Information Project CD-ROM. Alaska Department of Natural Resources. Anchorage, AK. 1992.

15. Kvenvolden, K. A.; Hostettler, F. D.; Carlson, P. R.; Rapp, J. B.; Threlkeld, C. N.; Warden, A. *Environ. Sci. Technol.* **1995**, *29*, 2684–2694.

16. Short, J. W.; Jackson, T. J.; Larsen, M. L.; Wade, T. L. In *Exxon Valdez Oil Spill Symposium*; Rice, S. D., Spies, R. B., Wolfe, D. A., Wright, B. A., Eds.; American Fisheries Society Symposium 18: Bethesda, Maryland,1996 ; pp 140–148.

17. Kvenvolden, K. A., Rapp, J. B, and Bourell, J. H. 1985. In *Alaska North Slope Oil/Rock Correlation Study*; Magoon, L. B., Claypool, G. E., Eds); American Association of Petroleum Geologists Studies in Geology, No. 20, 1985; pp. 593–617.

18. Short, J. W. *Environ. Forensics* **2002** *3*, 349–355.

19. Wooley, C. *Spill Sci. Technol.* **2002**, *7*, 89–104.

20. Wolfe, D. A.; Galt, J. A.; Watabayashi, G.; Short, J.; O'Claire, C.; Rice, S.; Michel, J.; Payne, J. R.; Braddock, J.; Hanna, S.; Sale, D. *Environ. Sci. Technol* **1994**, *28*, 561A–568A.

21. Bragg J. R.; Yang, S. H. In *Exxon Valdez Oil Spill: Fate and Effects in Alaskan Waters*; Wells, P. G., Butler, J. N., Hughes, J. S., Eds.; American Society for Testing and Materials Pub. 1219: Philadelphia, Pennsylvania, 1995; pp. 178–214.

22. Jahns, H. O.; Bragg, J. R.; Dash, L. C.; Owens, E. H. *Proceedings, 1991 International Oil Spill Conference*, American Petroleum Institute, Washington, D.C., 1991, pp. 167–176.

Tables

Table 1. Estimated beach area (ha) covered by surface oil within Prince William Sound, 12 years following the Exxon Valdez oil spill, based on SCAT surveys of randomly selected beaches. ND: not detected.

	Category I	Category II	Category III	Combined Categories
100 m Beach Segments	0.542	0.096	0.070	
<100 m Beach Segments	0.162	0.078	ND	
Combined Segments	0.704	0.174	0.070	0.948

Table 2. Estimated mean, coefficient of variation (%) and 95% confidence interval of beach area (ha) covered by surface oil within Prince William Sound, 12 years following the Exxon Valdez oil spill, based on stratified random sampling. ND: not detected.

	Category I	Category II	Category III	Combined Categories
100 m Beach Segments	1.21 (27.1%)	1.90 (65.7%)	ND	
	0.627 - 1.90	0.158 - 4.65		
<100 m Beach Segments	0.755 (32.4%)	0.264 (38.7%)	ND	
	0.327 - 1.27	0.0829 - 0.478		
Combined Segments	1.97 (21.2%)	2.16 (57.9%)	ND	4.13 (32.0%)
	1.19 - 2.81	0.386 - 4.94		2.07 - 7.705

Table 3. Estimated mean, coefficient of variation (%) and 95% confidence interval of beach area (ha) containing subsurface oil within Prince William Sound, 12 years following the Exxon Valdez oil spill, based on stratified random sampling. ND: not detected.

	Category I	Category II	Category III	Combined Categories
100 m Beach Segments	2.27 (28.6%)	2.90 (70.3%)	ND	
	1.14 - 3.67	0.238 - 7.14		
<100 m Beach Segments	1.29 (35.8%)	0.709 (59.3%)	0.640 (97.7%)	
	0.490 - 2.327	0.0913 - 1.66	0 - 1.99	
Combined Segments	3.56 (22.5%)	3.61 (57.7%)	0.640 (97.7%)	7.80 (29.2%)
	2.12 - 5.32	0.696 - 8.17	0 - 1.99	4.06 - 12.7

Table 4. Estimated mean, coefficient of variation (%) and 95% confidence interval of subsurface oil mass (metric tons, mt) within Prince William Sound, 12 years following the Exxon Valdez oil spill, based on stratified random sampling. ND: not detected.

	Category I	Category II	Category III	Combined Categories
100 m Beach	18.6 (28.6%)	17.5 (77.8%)	ND	
Segments	6.73 - 36.6	0.462 - 48.8		
<100 m Beach	7.53 (36.0%)	4.93 (66.2%)	7.01 (100.2%)	
Segments	3.01 - 13.8	0.305 - 12.4	0 - 24.1	
Combined	26.2 (31.6%)	22.4 (62.8%)	7.01 (100.2%)	55.6 (32.2%)
Segments	12.7 - 46.6	3.00 - 55.1	0 - 24.1	26.1 - 94.4

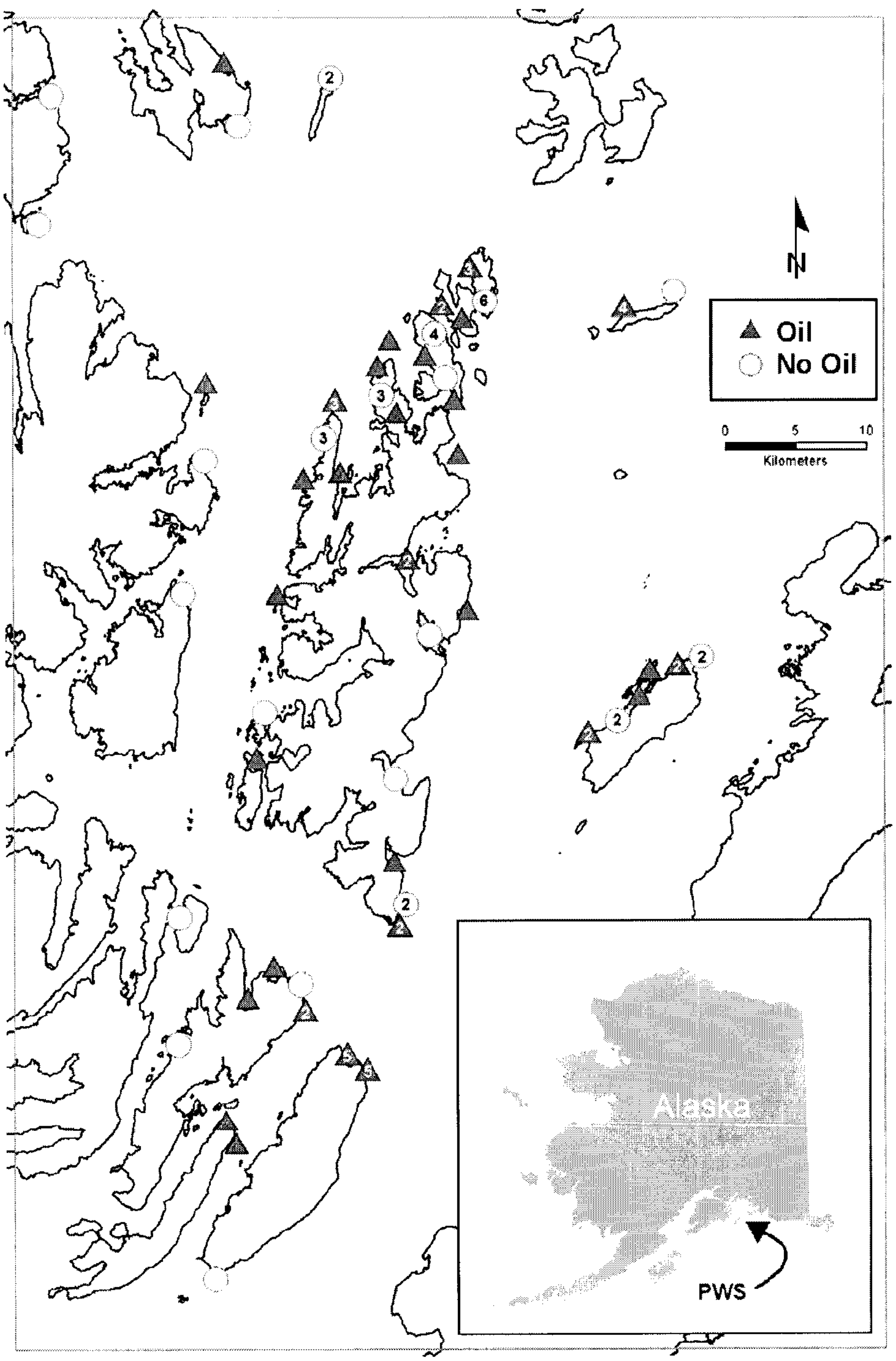


Figure 1. Map of the study area in Prince William Sound, Alaska. Symbols indicate stations where oil was detected (circles) or not (triangles) in our randomly placed quadrats. Numbers inside symbols indicate stations too close to resolve at this scale.

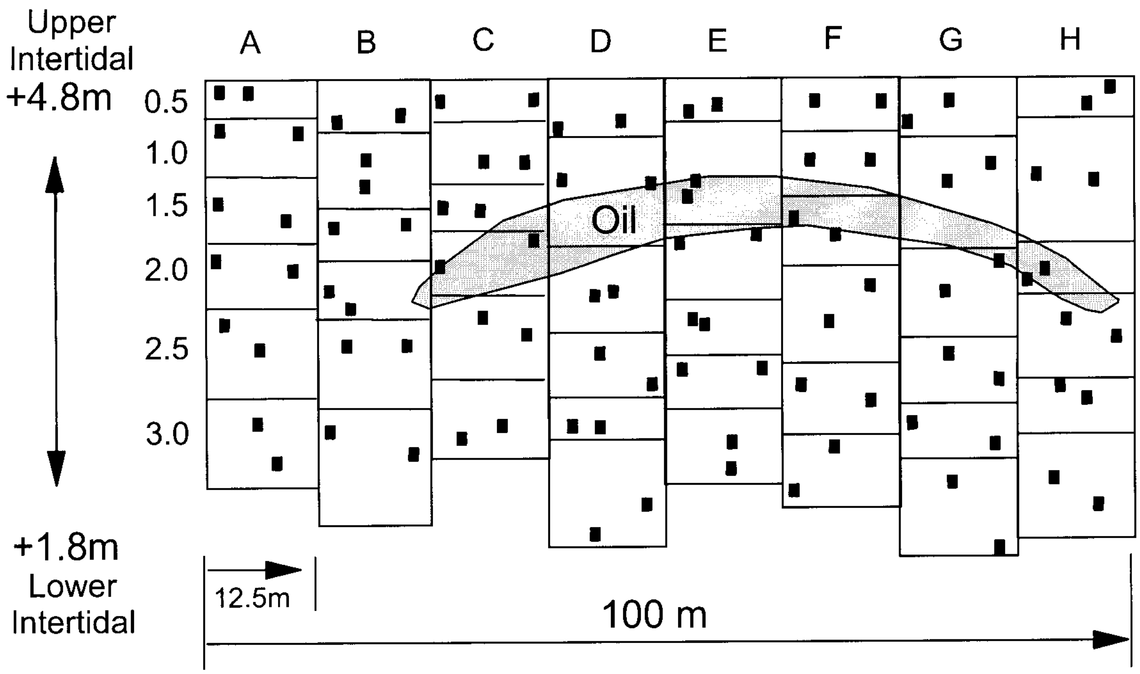


Figure 2. Diagram of a typical random sampling grid on a 100 m beach segment. The letters across the top denote 12.5 m columns, and the left-hand number column denotes 0.5 m vertical tidal drops from the upper margin of the sampling grid. The 96 black squares, two per 0.5 m vertical drop x column block, denote a typical random placement of quadrats within the sampling grid, and the grey-shaded area indicates a hypothetical oil patch. Random quadrat placements were made independently for each block and sampling grid. Sampled beach segments shorter than 100 m had fewer columns and randomly placed quadrats.

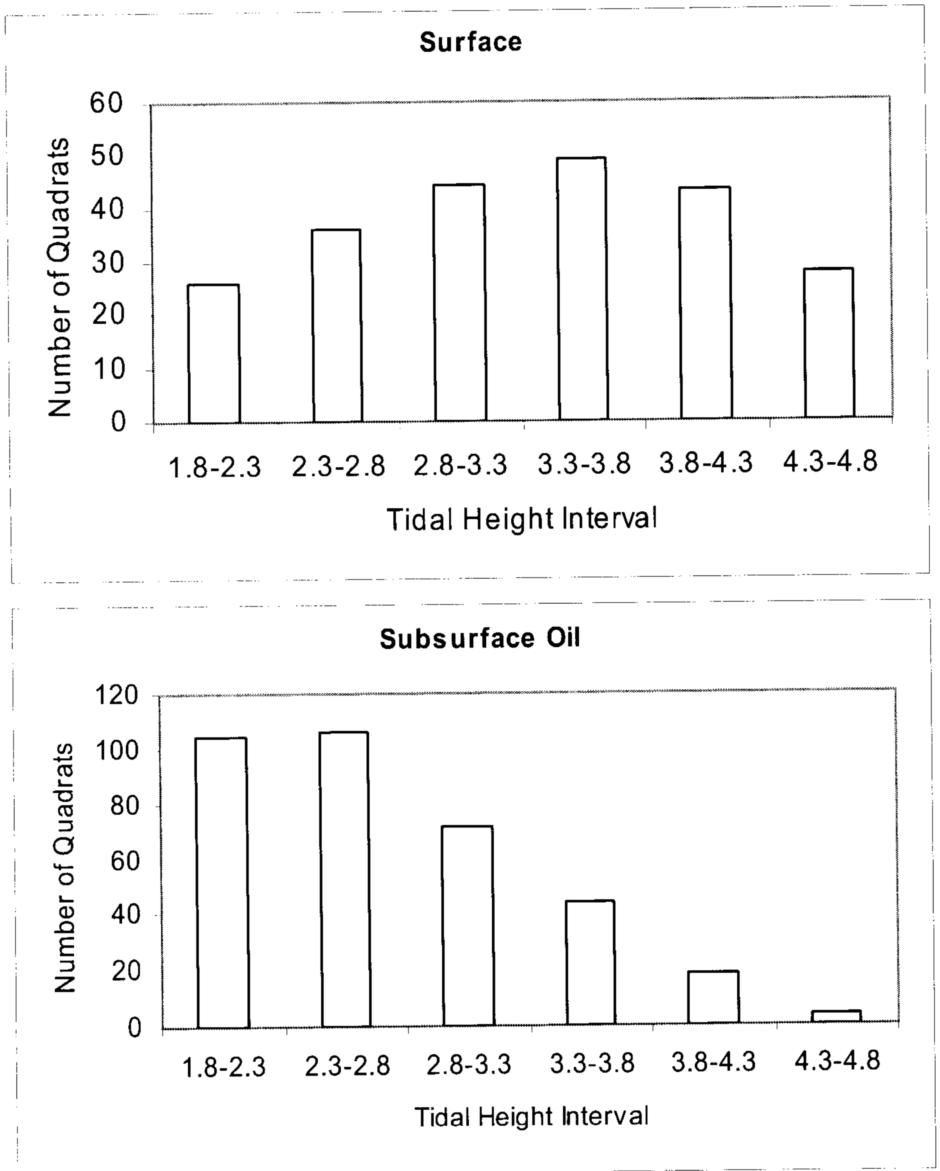


Figure 3. Distribution of surface (A) and subsurface (B) oil quadrats vs. tidal height.

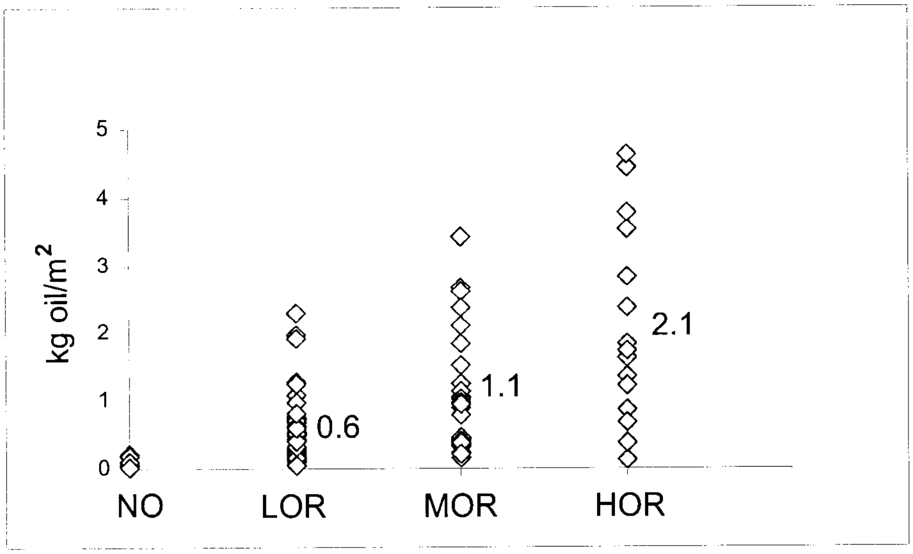


Figure 4. Oil mass per unit area for oil classification categories, LOR, MOR and HOR, which indicate light, medium and heavy oil residue, respectively (see Methods). Numbers plotted are category means.

Probability Sampling and Estimation of the Oil Remaining in 2001 from the 1989 *Exxon Valdez* Spill in Prince William Sound, Alaska

Jerome J. Pella* and Jacek M. Maselko

Auke Bay Laboratory, Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA, 11305 Glacier Highway, Juneau, Alaska 99801-8626

*Corresponding author telephone: 907-789-6027; fax: 907-789-6094; e-mail: Jerry.Pella@noaa.gov

Abstract

Disputes concerning the amount of spill oil remaining in Prince William Sound, Alaska, from the *Exxon Valdez* grounding in 1989 required revisiting the region in 2001. In contrast to earlier surveys based on purposeful selection of sampling locations, probability sampling was applied in order that unbiased estimates with measures of their precision could be computed. Beach segments and subsegments were stratified by their oiling histories and lengths, and random samples were selected for a visit. At each beach visited, the surface was grided into tidal elevation intervals and perpendicular columns, and every intersecting block was further subdivided into quadrats (0.25 m^2 areas) for random and adaptive sampling. Adaptive sampling pursued oil found in initial random quadrats in order to delimit entire patches. Subsurface oiled sediments were classified to a visual scale, and oil present in selected quadrats was extracted and weighed in a calibration study. The surface and subsurface oiled areas of the sediments at the visited beaches were estimated, together with the weight of oil in their subsurface sediments from the calibrated visual scale. Conservative estimates of oiled areas and weights for the visited beaches included only oil seen at random and adaptive quadrats, but unbiased estimates were computed by expansion for quadrats not sampled. The estimates at visited beaches were expanded for unsampled beach segments of strata, and summed for the total in Prince William Sound. Precision was evaluated by analytical formulas as well as by bootstrap resampling. Unbiased estimation of oiled areas and weights from oil found at the random and adaptive quadrats with precision evaluated by bootstrap resampling determined that until 2001, Prince William Sound still had a total of 41 thousand m^2 surface oiled area (95% interval, 20.7 m^2 - 70.5 m^2), and a total of 71 thousand m^2 subsurface oiled area (95% interval, 37.7 m^2 - 113.2 m^2), having subsurface oil weighing 50 m.t. (95% interval, 24.4 m.t. - 82.6 m.t.). Unbiased estimates based on the random quadrats only were 78 thousand m^2 for subsurface oiled area (95% interval, 40.6 m^2 - 127.3 m^2), and 56 m.t. (95% interval, 26.1 m.t. - 94.4 m.t.) for subsurface oil mass, agreeing well with the estimates based on combined random and adaptive quadrats when considered in light of sampling error..

Introduction

On March 24, 1989, the oil tanker *Exxon Valdez* ran aground on Bligh Reef in Prince William Sound, Alaska, releasing 41 million liters of Alaskan North Slope crude oil into the sea (Fig. 1). Subsequent linear estimates for the oil's dispersal on beaches declined from a beach distance of 783 km in 1989 to 10 km in 1992 (Neff et al, 1995), while the remaining oil volume in the fall of 1991 was estimated at 40,882 liters (Koons and Jahns, 1992). The traditional methodology to estimate volume of oil stranded on shorelines uses visual surveys together with average oil depths. The accuracy of both linear and volume estimates depends on the validity of an assumption that absence of surface oil on beaches indicates the same condition for subsurface oil (Finkelstein and Gundlach 1981; Owens, 1987). This supposition is plausible shortly following the stranding of oil on the beaches, but as time passes after the spill, surface oil becomes a poor indictor of subsurface oil.

Neff et al. (1995) reported that in 1991, about 2 years after the spill, only 33% of pits with surface oil also had subsurface oil, but importantly, no information was available regarding the occurrence of subsurface oil without surface presence. When residents of Prince William Sound complained of oil persistence at a local subsistence beach, a shoreline cleanup in 1996 and 1997 about 8 years after the spill unearthed substantial deposits of subsurface oil that were not evident from surface observations (Brodersen et al., 1999). During the present survey in 2001 about 12 years after the spill, only 24 of 225 (11%) pits with surface oil also had subsurface oil, and, more importantly, only 24 of 341 (7%) with subsurface oil also had surface oil.

Previous studies targeting subsurface oil have used purposeful selection of sampling locations (Gibeaut and Piper, 1998; Neff et al, 1995) in place of long-standing methods of probability sampling (e.g., see Cochran 1963). Both Gibeaut and Piper (1998) and Neff et al. (1995) selected shoreline sampling locations in Prince William Sound based on their oiling histories and recommendations from the public. Probability sampling to select locations for estimating remaining oil was not applied in either survey. In contrast to probability sampling, purposeful selection does not guarantee unbiasedness of estimation, nor does it allow an evaluation of accuracy and precision of estimates. Owens (1987), Neff et al. (1995), and Gibeaut et al. (1993) concede that sampling crew judgment has a large effect on such survey estimates, especially when applied to subsurface oil. Both Neff et al. (1995) and Gibeaut and Piper (1998) took a systematic sampling approach at the selected sites by delineating the extent of subsurface oil patches, similar to the adaptive sampling plan followed by the present Auke Bay Laboratory survey at randomly selected sites. Critical to probability-based estimation with adaptive sampling is a random selection of initial pits at a selected site, but both Neff et al. (1995) and Gibeaut and Piper (1998) failed to describe how locations of the initial subsurface pits at a site were chosen.

The use of purposeful selections of sampling units by the different surveys resulted in conflicting estimates of oil persistence. Whereas Neff et al (1995) report 12,000 m^2 area of subsurface oil remaining in 1992, Gibeaut and Piper (1998) conclude that in 1993, the remaining subsurface oil area in Prince William Sound was closer to 33,749 m^2. The 1993 survey also provided a volume estimate for subsurface oiled sediments adjacent to only the surveyed shoreline, equal to 2,041 m^3. Probable subsurface oiled sediments adjacent to unsurveyed shoreline were omitted. Neff et al. (1995) estimated oil to persist on 96 km in 1991 and 10 km in 1992. Because probability sampling was not used previously by either Exxon or public (Exxon Valdez Oil Spill Trustee Council) sponsored scientists and the actual amounts of oil remaining were unknown, of course, no basis occurred by which their accuracy or precision could be appraised. Therefore, the disagreements over the remaining amounts of oil among scientists required another survey for oil in Prince William Sound, this time using probability sampling so that accuracy and precision of the estimates could be evaluated. The Exxon Valdez Trustees Oil Spill Council contracted with the Auke Bay Laboratory in October, 2000, to perform the survey and estimation.

This manuscript describes the Auke Bay Laboratory survey sampling design, provides related estimation formulas, and reports the point estimates and their precision which are computed for the dispersal area and quantity of the remaining oil. The dispersal area is determined from a stratified random sampling design of beaches in Prince William Sound, based on their oiling histories. The oiled surface and subsurface areas of selected beaches are estimated by initial random sampling of smaller comprising units called quadrats, followed by delineation of discovered oil patches by systematic search. The quantity, or weight, of oil remaining is determined from visual observations of oiling at sampled quadrats in the survey, which are transformed to oil weights from data of a calibration study. Two approaches to evaluate the precision of estimation are compared—analytical formulas derived from sampling theory (Cochran 1963, Thompson 1992), and a computer-intensive method called bootstrap resampling (Efron 1982).

Methods

Sampling design

Segments and Subsegments

Following the 1989 Exxon Valdez oil spill, Prince William Sound beaches have been surveyed repeatedly for the prevalence of oil. The main surveys occurred in 1989, 1990, 1991 and 1993, and their goal was to map the geographical extent of the spill oil on shorelines classified to several contamination levels. The surveys measured subsurface contamination, or oiling, levels by a coarse ordered scale, whose decreasing values were termed “heavy”, “medium”, “light”, “oil film/trace” and “clean”. The surveys related uninterrupted horizontal distances of such contamination levels to a list of adjacent shoreline lengths, or beach segments. Three separate oiling categories of these beach segments and their total shoreline lengths were identified for which our expectations of finding oil were highest:

(1) Heavy oiling found during 1990, 1991 and 1993 surveys, 24.4 km.

(2) Medium oiling found 1990, 1991 and 1993 surveys, 49.1 km.
(3) Heavy oiling found in 1989 survey, but becoming less than heavy oiling more recently, 43.1 km.

Other shorelines of Prince William Sound that had medium and low oiling in 1989 and low oiling more recently were assumed to be clean without further measurement in 2001, and therefore were excluded from the estimation for total oil persisting in Prince William Sound.

The total shoreline within any oiling category consisted of beach segments of varying lengths. Longer beach segments were divided into 100 m lengths for sampling. This division was necessary because the most that field crews could sample between tides each day was about one hundred meters. Leftover subsegments less than 100 m resulted. Therefore, each oiling category contained two subcategories, 100 m and < 100 m beach subsegments. Overall, 6 strata of beach subsegments resulted, being the possible combinations of 3 oiling categories and 2 subsegment length categories.

Beach subsegments to be sampled were selected at random from the subcategories. For 100 m subsegments, simple random selection without replacement was used. For <100 m subsegments, random selection with probability proportional to length (ppl) was used. The ppl sampling requires replacement.

Within Subsegment Sampling

Because the oil was not deposited on the shorelines uniformly with respect to tidal elevation, the beach surface of each selected subsegment was divided into six 0.5 m tidal elevation intervals (0-0.5, 0.5-1.0, 1.0-1.5, 1.5-2.0, 2.0-2.5, 2.5-3.0, referring to meters below mean high tide). These intervals covered the beach between the highest elevations that deposition could have occurred and the lowest elevation accessible diurnally to surface sampling.

Stratified adaptive sampling (Thompson 1992) was used within subsegments because the oil was expected to be distributed in rare patches and mapping these patches would improve precision. Such delineation of the oil distribution was consistent with the earlier traditional survey methods, but at a much finer geographic scale. In order to distribute the sampling over the beach surface of a subsegment, the surface was divided into vertical, equal-width, columns. In the case of 100 m subsegments, 8 columns, 12.5 m wide, were formed. Shorter beach subsegments had fewer columns of about the same width. The intersections of horizontal elevation intervals and vertical columns on the beach surface resulted in blocks (e.g., 48 blocks per 100 m subsegment), each of which was subsampled for the presence of oil.

The surface of each block consisted of sampling units of 0.5 m by 0.5 m quadrats. Two sampling units per block were chosen at random without replacement. For a 100 m beach subsegment, a total of 96 quadrats were chosen. The surface of each randomly selected sampling unit was scrutinized for presence of oil, and next excavated to a maximum depth of 0.5 m, or (often) less if bedrock was present. The resulting pit is named an origin pit, to distinguish it from any succeeding pits, called adaptive pits. Any subsurface oil found in the origin pit was visually classified by increasing quantity: oil film (OF), light oil residue (LOR), medium oil residue (MOR), and heavy oil residue (HOR). Digging the origin pit could disturb the surface of nearby quadrats, and so surface oil observations were restricted to the origin quadrats. For subsurface oil, no confusion occurred between oil present before sampling and that recently disturbed by digging pits. If subsurface oil was found in the origin pit, a neighborhood of 4 bordering adaptive pits, each of the same dimensions as the origin pit, was dug about it–above, below, to the right and left. When oil was present in one or more of these adaptive pits, additional neighborhoods of adaptive pits were dug about each. Eventually, the entire patch of subsurface oil was uncovered, which could extend even beyond the block of the origin pit. The record for each randomly selected sampling unit indicated the presence or absence of surface oil, as well as the visual index of amount of subsurface oil, and the associated numbers of succeeding adaptive pits and their visual indices of subsurface oil, if any.

The information about oil from the pits was used to estimate the area occupied by subsurface oil per stratum by three modes–Stratified Adaptive (Adaptive), Stratified Random Sampling (SRS), and Observed. The Adaptive estimate for each stratum used both the randomly selected and adaptive pits, first expanding oiled area within sampled subsegments for the unsampled portion, and then expanding estimated oiled areas in sampled subsegments for unsampled subsegments in the stratum. The SRS estimate differs from the Adaptive estimate in that only the randomly selected origin pits were used, and the expansion within sampled subsegments differs as a result. Either the Adaptive or the SRS estimator of oiled area per stratum is theoretically unbiased. The Observed estimate used the oil patches found by the random and adaptive pits without expansion for the unsampled portion within subsegments, but with expansion for unsampled subsegments in the stratum. The Observed estimator of oiled area per stratum is biased low, and no variance is computed for uncertainty at the beach segment level. Because the adaptive pits were not scored for surface oil, neither an Adaptive estimate, nor an Observed estimate, was available for surface oil, only an SRS estimate based on the randomly selected origin pits. Because information about the weight of oil derived from subsurface observations, and not surface observations, only subsurface oil weight could be estimated using the Adaptive, SRS, and Observed approaches.

Gravimetric samples

Gravimetric samples of oiled beach material were collected in order to calibrate the visual oiling classification to the physical amounts of oil present in the pits. A total of 100 pits were sampled on a schedule of 0-5 pits per day throughout the field season in order to obtain broad coverage of Prince William Sound beaches for the four oiling categories (OF, LOR, MOR, HOR). Each gravimetric sample was a subsample of the thoroughly-mixed material from an oiled pit. The weight of oil present in each subsample was determined by chemical extraction and gravimetric measurement, and the total weight of oil in the pit was estimated by simple expansion for subsampling. The estimated total weight of oil per unit surface area for the pit was obtained from the pit surface area of 0.25 m^2. The oiling category OF was subsequently dropped from the analysis as only 2 gravimetric samples were taken in the field. Weight of oil for the OF category was assumed to equal zero.

Estimation Formulas

Oiled-area for a selected beach subsegment

A stratified estimator of modified Horvitz-Thompson type (chapter 26, Thompson, 1992) provides each subsegment estimate of oiled area, $\hat{y}$, and its variance, e.g.,

$$\hat{y} = \sum_{i=1}^{K} \frac{a_i I_i}{p_i}, \quad \text{var}(\hat{y}) = \sum_{i=1}^{K} \sum_{j=1}^{K} \frac{a_i a_j I_i I_j}{p_{ij}} \left(\frac{p_{ij}}{p_i p_j} - 1 \right),$$

where

K oil patches occur in the entire subsegment,
a_i is the area of the ith patch,
I_i is a 0-1 indicator for the intersection of the ith patch by the initial pairs of random quadrats from the L blocks comprising the surface of the beach subsegment,
p_i is the probability of this intersection, and
p_{ij} is the probability of intersection of both the ith and jth patches by the initial pairs of random quadrats.

The probability of inclusion of the ith patch in the initial random quadrats is

$$p_i = 1 - \prod_{k=1}^{L} \binom{N_k - x_{ki}}{2} \bigg/ \binom{N_k}{2},$$

where

N_k quadrats comprise the surface of the kth block, and
x_{ki} quadrats in the kth block intersect the ith patch.

The probability that the initial random quadrats intersect both the ith and jth patches is

$$p_{ij} = 1 - (1 - p_i) - (1 - p_j) + \prod_{k=1}^{L} \binom{N_k - x_{ki} - x_{kj}}{2} \bigg/ \binom{N_k}{2}.$$

Oiled-area estimation for strata

Prince William Sound beach subsegments were partitioned into 3 oiling categories and 2 length categories. Consider an oiling category in which N_1 subsegments were 100 m, and N_2, shorter lengths, denoted by $L_1, L_2, \ldots, L_{N2}$. Let n_1 100 m subsegments have been drawn at random without replacement from the N_1 available, and n_2 subsegments of the second category have been drawn with ppl sampling from the N_2 available. To do ppl sampling, n_2 units were drawn with replacement and probabilities $p_i = L_i/(L_1 + \ldots + L_{N2}) = L_i/L$. Adaptive stratified sampling (Thompson 1992) was applied to each selected beach subsegment for estimation of oiled area and oil weight. Let $y_{11}, y_{12}, \ldots, y_{1n1}$ be the unobserved oiled areas in the selected beach subsegments of the first stratum, and let $\hat{y}_{11}, \hat{y}_{12}, \ldots, \hat{y}_{1n1}$ be the corresponding estimated oiled areas. Let $y_{21}, y_{22}, \ldots, y_{2n2}$ and $\hat{y}_{21}, \hat{y}_{22}, \ldots, \hat{y}_{2n2}$ be the unobserved and estimated oiled areas for the ppl sample from the second stratum.

The unobserved total oiled area from the first stratum is $\sum_{i=1}^{N_1} y_{1i} = T_1$. If the y_{1i} were observable, an appropriate estimate of total oiled area would be $\hat{T}_1 = N_1 \overline{y_1}$, and its estimated variance would be $\text{var}(\hat{T}_1) = N_1(N_1 - n_1)\frac{s_1^2}{n_1}$, where s_1^2 is the sample variance of the y_{1i}'s (p. 103, Thompson 1992). Because the y_{1i}'s are not observable and must be estimated, the estimate becomes

$$\hat{T}_1 = N_1\tilde{y} = N_1\left(\frac{\sum_{i=1}^{n_1} \hat{y}_{1i}}{n_1}\right). \quad (1)$$

The corresponding estimated variance of $\hat{T}_1$ is

$$\hat{V}(\hat{T}_1) = N_1(N_1 - n_1)\frac{s_1^2}{n_1} + \frac{N_1}{n_1}\left(\sum_{i=1}^{n_1} \text{var}(\hat{y}_{1i})\right), \quad (2)$$

where s_1^2 is the sample variance of the $\hat{y}_{1i}$'s, and $\text{var}(\hat{y}_{1i})$ is the estimated error variance for the ith beach subsegment from stratified adaptive sampling. The first term of the estimated variance represents the variation among 100 m beach subsegments, and the second term is the contribution due to estimation error of the selected subsegments. The variance estimator is obtained from Equation 6 on p. 129 of Thompson (1992), but with $\text{var}(\hat{y}_{1i})$ from stratified adaptive sampling substituted for the variance of the estimated total oiled area had simple random sampling of quadrats within the beach subsegment been used. Thompson's Equation 6 is an unbiased estimate of the variance for a population total obtained by simple random sampling at each of two stages, primary units and secondary subunits (Also see Appendix I).

The unobserved total area from the second stratum is $\sum_{i=1}^{N_2} y_{2i} = T_2$. Were the y_{2i} observable and drawn with replacement and probability equal to L_i/L ($L = \Sigma L_i$), the Hansen-Hurwitz estimator (p. 47, Thompson 1992) for T_2 would be $\hat{T}_2 = \frac{L}{n_2}\sum_{i=1}^{n_2}\frac{y_{2i}}{L_i}$ with each value, y_{2i}, used in the sum as often as the ith beach subsegment was selected. The estimated variance of the Hansen-Hurwitz estimator would be

$$\hat{V}(\hat{T}_2) = \frac{L^2}{n_2(n_2 - 1)}\sum_{i=1}^{n_2}\left(\frac{y_{2i}}{L_i} - \frac{\hat{T}_2}{L}\right)^2.$$

Because the y_{2i}'s are not observable and must be estimated,

$$\hat{T}_2 = \frac{L}{n_2}\sum_{i=1}^{n_2}\frac{\hat{y}_{2i}}{L_i}. \tag{3}$$

The corresponding estimated variance of $\hat{T}_2$ is,

$$\hat{V}(\hat{T}_2) = \frac{L^2}{n_2(n_2-1)}\sum_{i=1}^{n_2}\left(\frac{\hat{y}_{2i}}{L_i} - \frac{\hat{T}_2}{L}\right)^2, \tag{4}$$

where values of the $\hat{y}_{2i}$'s correspond to the estimated total in the ith beach subsegment selected (repeated selections of a particular subsegment result in differing estimates included in the sum). This variance formula is described on p.132 of Thompson (1992).

Sampling for oiled-area estimation is done independently between subsegment-length strata and among the 3 oiling categories. Therefore, estimates of total oiled area in Prince William Sound is the sum of the 6 oiled-area estimates (3 oiling categories x 2 subsegment-length categories) and the variance of this sum is the sum of the 6 individual variances.

Estimates of oil weight are not done independently among subsegment-length strata and oiling categories, as we see next.

Oil-weight estimation

In order to estimate the weight of oil on the beaches, a discrete visual scale having M (= 4) levels of oil contamination was defined. Consider a particular subsegment-length stratum of an oiling category. Denote its total oiled areas belonging to the M visual categories of oil contamination by $\mathbf{A}'=(A_1,\ldots,A_M)$, and let μ_{wm} be the average weight of oil per unit area of A_m. Then the weight of oil on its beaches is the total of visual-category products of the unobserved oiled areas and average weights of oil per unit area, *viz.*,

$$W = \sum_{m=1}^{M} A_m \mu_{wm}.$$

Estimates of areas in a stratum belonging to these visual categories are derived from the field sampling data. The field sampling for each subsegment allowed estimation of both the total oiled subsurface area and the proportions of that area belonging to the visual categories. Recall that in regular field sampling, the beach surface of each selected 100 m subsegment was divided into 48 blocks, and each member of a pair of random origin pits from each block was examined for surface and subsurface oil. If the subsurface was oiled, the origin pit was classified to a visual category of contamination. The beach surface of shorter < 100 m subsegments was divided into fewer blocks, but again each member of a pair of random origin pits from each block was classified to a visual category of subsurface oiling. The numbers of origin pits in the oiled area of a sampled beach subsegment was random, depended on the size of the oiled area, and actually ranged from 1 to 31, out of a possible 96, for 100 m subsegments, and 1 to 35 for <100 m subsegments. Later, these subsegment counts that classified to the M visual categories are modeled with the multinomial probability function when estimating the visual-category proportions that compose the subsurface oiled area of a sampled beach subsegment.

Estimates of the average weight of oil per unit area within visual oiling categories were obtained from the gravimetric samples of material in oiled origin pits. Denote by $n_{w1}, n_{w2}, \ldots, n_{wM}$ the numbers of pits from each of the *M* visual categories that were processed for weight of oil. The underlying means and variances for the weight of oil per unit area in the *M* visual categories, $\mu_{w1}, \mu_{w2}, \ldots, \mu_{wM}$, and $\sigma_{w1}^2, \sigma_{w2}^2, \ldots, \sigma_{wM}^2$ are estimated from the gravimetric samples by sample averages of their weights per unit area, $\overline{w}_1, \overline{w}_2, \ldots, \overline{w}_M$, and associated sample variances, $s_{w1}^2, s_{w2}^2, \ldots, s_{wM}^2$, respectively.

The oil weight estimation will be described in reverse order for the two categories of beach subsegments, beginning with <100 m subsegments and finishing with 100 m subsegments.

<100 m beach subsegments

The total weight of spill oil in the second stratum comprised of <100 m beach subsegments within an oiling category will be denoted by W_2. This weight can be written as

$$W_2 = \sum_{m=1}^{M} A_m \mu_{wm},$$

where A_m is the oiled area comprised of visual category m, and μ_{wm} is the average weight of oil per unit area of A_m. Each sampled beach subsegment provides an estimate of $\mathbf{A}'=(A_1,\ldots,A_M)$, the visual-category areas of the entire stratum. If a single beach subsegment, say the i-th subsegment, were drawn, the Hansen-Hurwitz estimate of **A** would be

$$\hat{\mathbf{A}}_i = \begin{pmatrix} \frac{L}{L_i}\hat{y}_i\hat{p}_{1i} \\ \frac{L}{L_i}\hat{y}_i\hat{p}_{2i} \\ \vdots \\ \frac{L}{L_i}\hat{y}_i\hat{p}_{Mi} \end{pmatrix},$$

where $\hat{y}_i$ is the estimate of the oiled area from adaptive sampling, and $\hat{\mathbf{p}}_i = (\hat{p}_{1i},\ldots,\hat{p}_{Mi})'$ is the vector of estimated proportions of the oiled area belonging to the *M* visual categories in the beach subsegment. The beach subsegments were drawn independently with replacement, and if a beach subsegment was drawn more than once, its oiled area and area proportions of the *M* visual levels were estimated repeatedly and independently of previous draws. As a result, the n_2 estimates, $\hat{\mathbf{A}}_i$, $i=1,\ldots,n_2$, are independent and identically distributed random vectors from an underlying multivariate sampling probability distribution for estimated stratum totals of the *M* visual categories of oiled areas. Their average,

$$\overline{\mathbf{A}} = \begin{pmatrix} \overline{A}_1 \\ \overline{A}_2 \\ \vdots \\ \overline{A}_M \end{pmatrix} = \frac{1}{n_2}\sum_{i=1}^{n_2} \hat{\mathbf{A}}_i = \frac{1}{n_2}\sum_{i=1}^{n_2} \frac{L\hat{y}_i}{L_i}\begin{pmatrix} \hat{p}_{i1} \\ \hat{p}_{i2} \\ \vdots \\ \hat{p}_{iM} \end{pmatrix}, \quad (5)$$

is an unbiased estimate of the visual oiled areas in the stratum, and it has unbiased variance-covariance matrix estimate,

$$S_{\overline{A}} = (s_{\overline{A}_i,\overline{A}_j}) = \begin{pmatrix} s^2_{\overline{A}_1} & s_{\overline{A}_1,\overline{A}_2} & \cdots & s_{\overline{A}_1,\overline{A}_M} \\ - & s^2_{\overline{A}_2} & \cdots & s_{\overline{A}_2,\overline{A}_M} \\ - & - & \ddots & \vdots \\ - & - & - & s^2_{\overline{A}_M} \end{pmatrix} = \frac{1}{n_2(n_2-1)} \cdot$$

$$\begin{pmatrix} \sum_{i=1}^{n_2}\left(\hat{A}_{i1}-\overline{A}_1\right)^2 & \sum_{i=1}^{n_2}\left(\hat{A}_{i1}-\overline{A}_1\right)\left(\hat{A}_{i2}-\overline{A}_2\right) & \ldots\sum_{i=1}^{n_2}\left(\hat{A}_{i1}-\overline{A}_1\right)\left(\hat{A}_{iM}-\overline{A}_M\right) \\ ------ & \sum_{i=1}^{n_2}\left(\hat{A}_{i2}-\overline{A}_2\right)^2 & \ldots\sum_{i=1}^{n_2}\left(\hat{A}_{i2}-\overline{A}_2\right)\left(\hat{A}_{iM}-\overline{A}_M\right) \\ \vdots & \vdots & \vdots \\ ------ & ----- & \ldots\sum_{i=1}^{n_2}\left(\hat{A}_{iM}-\overline{A}_M\right)^2 \end{pmatrix}. \quad (6)$$

The unknown weight of oil in the stratum, W_2, is estimated as

$$\hat{W}_2 = \sum_{m=1}^{M} \overline{A}_m \overline{w}_m, \quad (7)$$

and it has variance (See Appendix II)

$$V(\hat{W}_2) = \sum_{i=1}^{M}\left\{A_i^2\frac{\sigma^2_{wi}}{n_{wi}} + \mu^2_{wi}V(\overline{A}_i) + V(\overline{A}_i)\frac{\sigma^2_{wi}}{n_{wi}}\right\} + 2\sum_{i=1}^{M}\sum_{j<i}\mu_{wi}\mu_{wj}Cov(\overline{A}_i,\overline{A}_j)$$

.

This variance is estimated unbiasedly by

$$\hat{V}(\hat{W}_2) = \sum_{i=1}^{M}\left\{\overline{A}_i^2\frac{s^2_{wi}}{n_{wi}} + \overline{w}_i^2 s^2_{\overline{A}_i} + s^2_{\overline{A}_i}\frac{s^2_{wi}}{n_{wi}}\right\} + 2\sum_{i=1}^{M}\sum_{j<i}\overline{w}_i\overline{w}_j s_{\overline{A}_i,\overline{A}_j}, \quad (8)$$

where $s_{\overline{A}_i}{}^2$ the estimated variance of $\overline{A}_i$, and $s_{\overline{A}_i,\overline{A}_j}$ is the estimated covariance between $\overline{A}_i$ and $\overline{A}_j$.

100 m beach subsegments

The total weight of spill oil in the first subsegment-length stratum, comprised of 100 m beach subsegments, will be denoted by W_1. Let $\mathbf{a}_i=(a_{i1},a_{i2},...,a_{iM})'$ denote the oiled areas composing the visual categories in beach subsegment i, and $\mathbf{A}'=(A_1,...,A_M)$ denote the total areas by visual category in the entire stratum. For a sample of n_1 beach subsegments, the M-dimensional visual area arrays, $\mathbf{a}_1$, $\mathbf{a}_2$, ...,$\mathbf{a}_{n1}$ are estimated by adaptive sampling. The visual area estimates are denoted by the vectors, $\hat{\boldsymbol{a}}_1, \hat{\boldsymbol{a}}_2, \ldots, \hat{\boldsymbol{a}}_{n_1}$, where $\hat{\mathbf{a}}_i = \hat{y}_i \hat{\mathbf{p}}_i$. Recall that $\hat{y}_i$ is the estimated total oiled area from adaptive sampling in the ith beach subsegment, and $\hat{\mathbf{p}}_i$ the associated vector of estimated proportions of the oiled area belonging to the *M* visual categories. The visual-category areas in the stratum are estimated by expanding the sample average as

$$\hat{\boldsymbol{A}} = N_1 \cdot \left(\frac{\sum_{i=1}^{n_1} \hat{a}_{i1}}{n_1}, \frac{\sum_{i=1}^{n_1} \hat{a}_{i2}}{n_1}, \ldots, \frac{\sum_{i=1}^{n_1} \hat{a}_{iM}}{n_1} \right)' = N_1 \tilde{\boldsymbol{a}}' \qquad (9)$$

If the areas within each beach subsegment could have been directly observed, rather than estimated as was necessary, the variance-covariance matrix of these area estimates would be

$$V_{\hat{A}} = N_1 (N_1 - n_1) \frac{V_{\mathbf{a}}}{n_1},$$

where $V_{\mathbf{a}}$ is the population variance-covariance matrix of the arrays, $\mathbf{a}_1$, $\mathbf{a}_2$, ...,$\mathbf{a}_{N1}$. The estimate of this matrix would be

$$\hat{V}_{\hat{\mathbf{a}}} = N_1 (N_1 - n_1) \frac{S_{\mathbf{a}}}{n_1},$$

where $S_{\mathbf{a}}$ is the sample variance-covariance matrix with elements

$$s_{\mathbf{a},i,j} = \frac{1}{n_1 - 1} \sum_{k=1}^{n_1} \left(a_{ki} - \bar{a}_i \right) \left(a_{kj} - \bar{a}_j \right).$$

Here $\bar{a}_i = \frac{1}{n_1} \sum_{i=1}^{n_1} a_{ki}$ is the sample average area of visual category i.

Because the $\mathbf{a}_i$ were not observed and had to be estimated, the variance-covariance matrix of $\hat{\mathbf{A}}$ includes a term for this estimation error and is

$$V_{\hat{A}} = N_1 (N_1 - n_1) \frac{V_{\mathbf{a}}}{n_1} + \frac{N_1}{n_1} \sum_{k=1}^{N_1} V_{\hat{\mathbf{a}}_k - \mathbf{a}_k}.$$

Here $V_{\hat{\mathbf{a}}_k - \mathbf{a}_k}$ is the variance-covariance matrix of estimation errors in beach subsegment k (see Appendix III),

$$\mathbf{V}_{\hat{\mathbf{a}}_k - \mathbf{a}_k} = \left(y_k^2 + \sigma^2_{\hat{\mathbf{y}}_k}\right) \cdot V_{\hat{\mathbf{p}}_k - \mathbf{p}_k} + \sigma^2_{\hat{\mathbf{y}}_k}\left(\mathbf{p}_k \mathbf{p}_k{}'\right).$$

The matrix, $V_{\hat{p}_k - p_k}$, is the variance-covariance matrix of the estimated visual category proportions in the kth beach subsegment from multinomial sampling of origin pits in the oiled area.

The jth diagonal element of the matrix, $V_{\hat{a}_k - a_k}$, can be written as

$$\mathrm{v}_{\hat{a}_k - a_k, j, j} = \left(y_k^2 + \sigma^2_{\hat{y}_k}\right) \cdot \frac{p_{kj}(1 - p_{kj})}{h_k} + \sigma^2_{\hat{y}_k} p_{kj}^2,$$

where

p_{kj} is the proportion of the oiled area in the kth beach subsegment belonging to the jth visual category of oiling,
y_k is the oiled area of subsegment k,
$\sigma^2_{\hat{y}_k}$ is the variance of the estimated oiled area in subsegment k, and
h_k is the number of origin pits that were found to be oiled and were classified to the several visual categories.

The off-diagonal element in the ith row and jth column ($i \neq j$) is

$$\mathrm{v}_{\hat{a}_k - a_k, i, j} = -\left(y_k^2 + \sigma^2_{\hat{y}_k}\right) \cdot \frac{p_{ki} p_{kj}}{h_k} + \sigma^2_{\hat{\mathbf{y}}_k} p_{ki} p_{kj}.$$

An unbiased estimate of $V_{\hat{a}_k - a_k}$ is denoted by $S_{\hat{a}_k - a_k}$ with diagonal elements

$$s_{\hat{a}_k - a_k, j, j} = \left[\frac{\hat{y}_k^2}{h_k}\right]\left[\left(1 - \frac{1}{h_k}\right)^{-1} \hat{p}_{kj}(1 - \hat{p}_{kj})\right] + \hat{\sigma}^2_{\hat{y}_k}\left[\hat{p}_{kj} - \left(1 - \frac{1}{h_k}\right)^{-1} \hat{p}_{kj}(1 - \hat{p}_{kj})\right],$$

and off-diagonal elements,

$$s_{\hat{a}_k - a_k, i, j} = -\left[\frac{\hat{y}_k^2}{h_k} + \hat{\sigma}^2_{\hat{y}_k}\right]\left[\left(1 - \frac{1}{h_k}\right)^{-1} \hat{p}_{ki} \hat{p}_{kj}\right].$$

The estimate of $V_{\hat{\mathbf{A}}}$ is

$$S_{\hat{\mathbf{A}}} = N_1(N_1 - n_1)\frac{S_{\hat{a}}}{n_1} + \frac{N_1}{n_1}\sum_{\mathbf{k}=1}^{\mathbf{n}_1} S_{\hat{\mathbf{a}}_k - \mathbf{a}_k}, \tag{10}$$

where

$S_{\hat{\mathbf{a}}}$ is the sample variance-covariance matrix of the $\hat{\mathbf{a}}_k$, k=1 ,..., n_1, and

$S_{\hat{\mathbf{a}}_k - \mathbf{a}_k}$ is the estimated variance-covariance matrix of estimation errors in the kth sampled subsegment.

The unknown weight of oil in the stratum, W_1, is estimated as

$$\hat{W}_1 = \sum_{m=1}^{M} \hat{A}_m \overline{w}_m , \qquad (11)$$

and it has a variance of

$$V(\hat{W}_1) = \sum_{i=1}^{M} \left\{ A_i^2 \frac{\sigma_{wi}^2}{n_{wi}} + \mu_{wi}^2 V(\overline{A}_i) + V(\overline{A}_i) \frac{\sigma_{wi}^2}{n_{wi}} \right\} + 2\sum_{i=1}^{M} \sum_{j<i} \mu_{wi}\mu_{wj} Cov(\overline{A}_i, \overline{A}_j)$$

.

This variance is estimated unbiasedly by

$$\hat{V}(\hat{W}_1) = \sum_{i=1}^{M} \left\{ \hat{A}_i^2 \frac{s_{wi}^2}{n_{wi}} + \left(\overline{w}_i^2 - \frac{s_{wi}^2}{n_{wi}} \right) s_{\overline{A}_i}^2 \right\} + 2\sum_{i=1}^{M} \sum_{j<i} \overline{w}_i \overline{w}_j s_{\overline{A}_i, \overline{A}_j} , \qquad (12)$$

where $s_{\hat{A}_i}^2$ the estimated variance of $\hat{A}_i$, and $s_{\hat{A}_i, \hat{A}_j}$ is the estimated covariance between $\hat{A}_i$ and $\hat{A}_j$.

Total weight of spill oil remaining in Prince William Sound

Denote by A_{ijm} the total oiled area of visual category m in the jth subsegment-length stratum (j = 1, or 100 m, j = 2, or < 100 m) of the ith oiling category (i = 1, 2, or 3). The total subsurface oiled area in Prince William Sound by visual category equals the sum of individual estimates, viz.,

$$A_{m\bullet\bullet} = \sum_{i=1}^{3} \sum_{j=1}^{2} A_{ijm}, \quad m = 1,2,\ldots, M,$$

and the total weight of oil remaining in Prince William Sound is

$$W = \sum_{m=1}^{M} A_{m\bullet\bullet} \mu_{wm} .$$

The visual category totals, $A_{m\bullet\bullet}$, are estimated by replacing each A_{ijm} by its estimate (Equations 5 and 9) in this equation, and the total weight of oil remaining in Prince William Sound is estimated by

$$\hat{W} = \sum_{m=1}^{M} \hat{A}_{m\bullet\bullet} \overline{w}_m . \qquad (13)$$

The variance of the total weight of oil remaining is

$$V(\hat{W})=\sum_{m=1}^{M}\left\{A_{m\bullet\bullet}^{2}\frac{\sigma_{wm}^{2}}{n_{wm}}+\mu_{wm}^{2}V(\hat{A}_{m\bullet\bullet})+V(\hat{A}_{m\bullet\bullet})\frac{\sigma_{wm}^{2}}{n_{wm}}\right\}+2\sum_{m=1}^{M}\sum_{j<m}\mu_{wm}\mu_{wj}Cov(\hat{A}_{m\bullet\bullet},\hat{A}_{j\bullet\bullet})$$

This variance is estimated unbiasedly by

$$\hat{V}(\hat{W})=\sum_{m=1}^{M}\left\{\hat{A}_{m\bullet\bullet}^{2}\frac{s_{wm}^{2}}{n_{wm}}+\left(\overline{w}_{m}^{2}-\frac{s_{wm}^{2}}{n_{wm}}\right)\breve{s}_{\hat{A}_{m\bullet\bullet}}^{2}\right\}+2\sum_{m=1}^{M}\sum_{j<m}\overline{w}_{m}\overline{w}_{j}\breve{s}_{\hat{A}_{m\bullet\bullet},\hat{A}_{j\bullet\bullet}}, \qquad (14)$$

where

$$\breve{S}_{\hat{A}}=\begin{pmatrix}\breve{s}_{\hat{A}_{1\bullet\bullet}}^{2} & \breve{s}_{\hat{A}_{1\bullet\bullet},\hat{A}_{2\bullet\bullet}} & \cdots & \breve{s}_{\hat{A}_{1\bullet\bullet},\hat{A}_{M\bullet\bullet}}\\ \breve{s}_{\hat{A}_{2\bullet\bullet},\hat{A}_{1\bullet\bullet}} & \breve{s}_{\hat{A}_{2}}^{2} & \cdots & \breve{s}_{\hat{A}_{2\bullet\bullet},\hat{A}_{M\bullet\bullet}}\\ \vdots & & \ddots & \\ \breve{s}_{\hat{A}_{M\bullet\bullet},\hat{A}_{1\bullet\bullet}} & & \cdots\cdots & \breve{s}_{\hat{A}_{M\bullet\bullet}}^{2}\end{pmatrix}$$

is the estimated variance-covariance matrix for the visual area array comprised of elements, $\hat{A}_{m\bullet\bullet}, m=1,\ldots,M$. The elements of this covariance array are obtained by summing corresponding elements of the estimated variance-covariance matrices of visual areas, $S_{\overline{A}}$ **or** $S_{\hat{A}}$ (Equations 6 or 10), of the six combinations of subsegment-length and oiling categories.

The formulas for estimates and their standard errors (square root of variances) for oiled areas and weight of oil remaining in Prince William Sound are complete. Approximate 95% confidence intervals can be computed for any estimate by subtracting (lower limit) and adding (upper limit) 2 standard errors to the point estimate. If the distribution of the estimator is not approximately normal, confidence intervals are better computed by the bootstrap method. Bootstrap estimates of standard errors and confidence intervals are computed as described next.

Bootstrapping for determination of precision of estimates

Adaptive, SRS, and Observed estimates of oiled areas and weights are computed by bootstrapping methods described next. As before, the estimation will be described in reverse order for the two categories of beach subsegments, beginning with <100 m subsegments and finishing with 100 m subsegments The average, standard deviation, and lower 2.5 percentile and upper 97.5 percentile of the empirical distributions are tabulated and reported. The percentiles provide a 95% confidence interval.

1. Oiled Area, <100 m subsegments

The total oiled area in each of the three oiling categories (heavy oiling, 1990-1993; medium oiling, 1990-1993; and heavy oiling, 1989) has been estimated by the Hansen-Hurwitz formula from oiled areas, Observed, SRS, or Adaptive, found in n_2 ppl randomly selected beach subsegments of N_2 available (n_2 and N_2 vary among oiling categories). Let

$$z_i = \frac{L}{L_i}\hat{y}_i, \quad i = 1,2,\ldots,n_2,$$

denote the estimate of the stratum total oiled-area from the ith subsegment. The Hansen-Hurwitz estimate is the average of the z_i. The z_i of each stratum are independent and identically distributed random variables, so a standard bootstrapping method is used to draw samples by which to estimate precision of the estimated total oiled area for a stratum: 1,000 resamples of size n_2 are drawn with replacement from the n_2 original z-values of the sample of beach subsegments. Denote these values for the b-th resample as $z^*_{b1}, z^*_{b2}, \ldots, z^*_{bn_2}$. Notice that each resample is composed of the original z-values of the sampled subsegments, each repeated a random number of times from 0 to n_2. The stratum total oiled-area estimates computed by the Hansen-Hurwitz formula from subsegments of the b-th resample, say T_{2b}^*, $i=1,2,\ldots,1000$, equal the resample averages of the z^*-values. The average, standard deviation, and lower 2.5 percentile and upper 97.5 percentile of the empirical distribution of the 1000 T_{2b}^* are recorded for Adaptive, SRS, and Observed oiled areas.

2. Oiled Area, 100 m subsegments

The total oiled area in each of three strata (heavy oiling, 1990-1993; medium oiling, 1990-1993; and heavy oiling, 1989) has been estimated by simple expansion of average oiled areas among the n_1 sampled subsegments, either Observed, SRS, or Adaptive. To evaluate the precision of each stratum estimate of total oiled area, 1,000 bootstrap resamples of n_1 subsegments are drawn. The populations of subsegments and sampling method for generating these bootstrap samples depends on the stratum sampling fraction, n_1/N_1. If n_1/N_1 is small (<0.05), dependence caused by sampling without replacement from the finite population of size N_1 is ignored, and standard bootstrapping methods for independent and identically distributed random variables apply. The bootstrap samples are drawn with replacement from the n_1 subsegments originally sampled. If n_1/N_1 is larger ($\geq$0.05), the dependence is not ignored, and bootstrapping methods for finite populations apply. The method used approximates that suggested by Gross (1980), as described by Booth et al. (1994). The bootstrap samples are drawn without replacement from a population of $[N_1/n_1]$ +1 copies of the original n_1 subsegments, where the expression, $[\cdot]$, denotes the integer part of the argument. Notice that any resample is composed of the estimated oiled areas of the original sampled subsegments, each occurring between 0 and n_1 times if n_1/N_1 is small, or between 0 and $[N_1/n_1]$ +1 times if n_1/N_1 is larger . Denote the estimated oiled areas of the subsegments in the b-th bootstrap sample as $\hat{y}^*_{b1}, \hat{y}^*_{b2}, \ldots, \hat{y}^*_{bn_1}$. Let T_{1b}^*, $b=1,2,\ldots,1000$, be the stratum total oiled-area estimates computed by expansion of the average oiled areas among subsegments the ith resample, viz.,

$$T_{1b}^{*} = N_1 \tilde{y}_b^{*} = N_1 \left(\frac{\sum_{i=1}^{n_1} \hat{y}_{bi}^{*}}{n_1} \right).$$

The average, standard deviation, and lower 2.5 percentile and upper 97.5 percentile of the empirical distribution of the 1000 T_{1b}^{*} are recorded for Adaptive, SRS, and Observed oiled areas.

3. Visual-category composition of oiled areas, < 100 m or 100 m subsegments

Each subsegment sampled during the survey has an estimated visual-category composition, $\hat{\mathbf{p}}$, from the h random quadrats in its oiled area. Specifically, if the counts of quadrats in the four visual categories were $\mathbf{h}=(h_1,h_2,h_3,h_4)$, the visual-category composition of the oiled area is estimated as

$$\hat{\mathbf{p}} = \left(h_1 / h, \ldots, h_4 / h\right), \quad h = \sum_{i=1}^{4} h_i .$$

A total of 1,000 bootstrap samples of $\mathbf{h}=(h_1,h_2,h_3,h_4)$, are obtained by one of two procedures, depending on the apparent sampling fraction. If the sampling fraction, $h/\hat{y}$, is less than 0.05, a sample of h quadrats is drawn with replacement from the original h quadrats observed. If the sampling fraction exceeds 0.05, a sample of size h is drawn without replacement from a population composed of $[\hat{y}/h]+1$ copies of the original h random quadrats in the oiled area, where $[\cdot]$ truncates the number to its integer part. In either case, denote the bootstrap sample by $\mathbf{h}^{*}=(h_1^{*},\ldots,h_4^{*})$. Then the corresponding estimated visual-category composition is

$$\hat{\mathbf{p}}^{*} = (\hat{p}_1^{*}, \hat{p}_2^{*}, \ldots, \hat{p}_4^{*}) = \left(h_1^{*} / h, \ldots, h_4^{*} / h\right).$$

4. Average weight of oil per quadrat of the visual categories

The n gravimetric-extraction samples of visual category m are sampled with replacement n times to provide averages, $\overline{w}_{bm}^{*}, b = 1,2,\ldots,1000; m = 1,2,3,4$.

5. Oil weight for strata with subsegments <100 m

The algorithm for computing bootstrapped values of the total oil weight for an oiled category of <100 m subsegments is as follows:

1. Set $b=1$

2. Let $z_{b1}^{*}, z_{b2}^{*}, \ldots, z_{bn_2}^{*}$ denote the bth bootstrap resample of z-values for a stratum of <100 m subsegments obtained at Section 1. Corresponding to these z^{*}-values are the particular beach subsegments from which they were computed. For these subsegments, generate bootstrap samples for the visual-category compositions, $\hat{\mathbf{p}}_{b1}^{*}, \ldots, \hat{\mathbf{p}}_{bn_2}^{*}$ as described in Section 3 above.

3. Compute the estimated visual-category oiled areas in the stratum by

$$\overline{A}_b^* = \begin{pmatrix} \overline{A}_{b,1}^* \\ \vdots \\ \overline{A}_{b,4}^* \end{pmatrix} = \frac{1}{n_2}\sum_{i=1}^{n} z_{bi}^* \begin{pmatrix} \hat{p}_{bi,1}^* \\ \vdots \\ \hat{p}_{bi,4}^* \end{pmatrix}.$$

4. Compute the total weight of oil in the stratum by $\hat{W}_{2,b}^* = \sum_{m=1}^{4} \overline{A}_{b,m}^* \overline{w}_{b,m}^*$ using the bootstrapped samples of average weight per quadrat at Section 4.

5. If b < 1000, set b=b+1 and go to step 1. Otherwise, stop bootstrap sampling.

6. Oil weight for strata with subsegments of 100 m

The algorithm for computing bootstrapped values of the total oil weight for an oiled category of 100 m subsegments is as follows:

1. Set b =1

2. Let $\hat{y}_{b1}^*, \hat{y}_{b2}^*, \ldots, \hat{y}_{bn_1}^*$ denote the *b*-th bootstrap resample of oiled-area estimates for a stratum of 100 m subsegments obtained under Section 2. Corresponding to these oiled-area estimates are the particular beach subsegments from which they were computed. For these subsegments, generate bootstrap samples for the visual-category compositions, $\hat{\mathbf{p}}_{b1}^*, \ldots, \hat{\mathbf{p}}_{bn}^*$ as described above in Section 3.

3. Compute the visual-category areas of each beach subsegment in the bth bootstrap resample as

$$\hat{\boldsymbol{a}}_{bi}^* = \begin{pmatrix} \hat{a}_{bi1}^* \\ \vdots \\ \hat{a}_{bi4}^* \end{pmatrix} = \hat{y}_{bi}^* \begin{pmatrix} \hat{p}_{bi1}^* \\ \vdots \\ \hat{p}_{bi4}^* \end{pmatrix}, \; i=1,2,\ldots,n_1$$

4. Compute the estimate of total oiled area by visual-category in the stratum for the *b*-th bootstrap sample as

$$\hat{A}_b^* = \begin{pmatrix} \hat{A}_{b1}^* \\ \vdots \\ \hat{A}_{b4}^* \end{pmatrix}' = \frac{N_1}{n_1} \begin{pmatrix} \sum_{i=1}^{n_1} \hat{a}_{bi1}^* \\ \vdots \\ \sum_{i=1}^{n_1} \hat{a}_{bi4}^* \end{pmatrix}$$

5. Compute the total weight of oil in the stratum by $\hat{W}_{1,b}^* = \sum_{m=1}^{4} \hat{A}_{b,m}^* \overline{w}_{b,m}^*$ using the bootstrapped samples of average weight per quadrat at Section 4.

6. If b < 1000, set b=b+1 and go to step 1. Otherwise, stop bootstrap sampling.

The recipes above for total oiled area and oil weight apply to a single stratum. For Prince William Sound totals of area estimates, the *b*-th area estimates, T^*_{1b} and T^*_{2b}, are summed over the strata to provide the *b*-th bootstrap estimate of the Prince William Sound total oiled-area. For Prince William Sound totals of weight estimates, the *b*-th estimates, $\hat{W}_{1,b}^*$ and $\hat{W}_{2,b}^*$, are summed over the strata to provide the *b*-th bootstrap estimate of the Prince William Sound total oil weight.

Results and Discussion

Sampling was concentrated (17% - 20% sampling fraction, by number of subsegments, or 20% - 27% by length) on beaches for which the most recent surveys found heavy oiling (lines 1 and 2, Table 1). Lower intensity (3% - 5% sampling fraction, by number of subsegments, or 3% - 8%, by length) was directed to beaches found to have medium oiling during these surveys (lines 3 and 4, Table 2), and yet lower intensity (2% sampling fraction, by number of subsegments, and 2%, by length) was applied to beaches found with heavily oiling in earlier surveys, but less than medium oiling more recently. This effort distribution is recognized as less than ideal now that better information on oil distribution in Prince William Sound is available. However, the program was constrained by government and landowner permits to the initial design.

Comparison between subsurface oiled area estimates within sampled subsegments (strata combined) for the SRS and Adaptive modes (Fig 2) shows good agreement as expected, but with several material discrepancies for subsegments with larger oiled areas. In particular, SRS is known to have underestimated the oiled areas on four subsegments (Fig 3), a result not possible by adaptive estimation (Fig 4). In general, the estimated standard errors of oiled areas in the sampled subsegments reflect the reduced uncertainty by including the adaptive pits (Fig 5).

Estimates of the total subsurface oiled area per stratum and measures of their precision have been computed for Adaptive, stratified random sampling (SRS), and Observed modes, either from analytical formulas or by bootstrap resampling (Table 2). As expected, Observed estimates of oiled quadrats within subsegments, when extrapolated to stratum totals, provide minimal, biased values. Adaptive and SRS estimates are in good agreement for the more heavily sampled strata, 1 and 2, and in fair agreement as sampling intensity declines in strata 3, 4, and 5. Analytical and bootstrap computations are generally in good agreement, but only the bootstrap confidence intervals appropriately shift to accommodate the skew distributions. Considering the bootstrap computations for the Adaptive estimates as most trustworthy, strata 1 and 2 had considerable oiled area, 13 thousand m^2 (4.4 thousand m^2 - 24.6 thousand m^2), and 22 thousand m^2 (10.3 thousand m^2 - 36.0 thousand m^2), respectively. Stratum 4 may also have had substantial oiled area, estimated imprecisely at 27 thousand m^2 (2.4 thousand m^2 - 63.9 thousand m^2). The other strata, 3, 5, and 6, had less oiled area, also imprecisely determined as 6 thousand m^2 (0.9 thousand m^2 - 13.8 thousand m^2), 3 thousand m^2 (0 m^2 - 9.7 thousand m^2), and 0 m^2 (none observed), respectively. Although the separate strata provide relatively imprecise subsurface oiled area estimates, the grand sum for Prince William Sound is reasonably precise (Table 3), with Adaptive estimates from the bootstrap computations equaling 71 thousand m^2 (37.7 thousand m^2 - 113.2 thousand m^2). The differences in grand sum of subsurface oiled area for Prince William Sound, whether Adaptive or SRS estimates, and analytical or bootstrap computations, are relatively less concerning as well (Table 3).

Estimates of the total surface oiled area per stratum and measures of their precision have been computed for stratified random sampling (SRS), either from analytical formulas or by bootstrap resampling (Table 4). Analytical and bootstrap computations are generally in good agreement, but only the bootstrap confidence intervals accommodate the skew distributions reasonably. Considering the bootstrap computations as most trustworthy, the total oiled surface area was estimated for strata 1, 2, and 3 as 8 thousand m^2 (3.3 thousand m^2 - 12.7 thousand m^2), 12 thousand m^2 (6.3 thousand m^2 - 19.0 thousand m^2), and 3 thousand m^2 (0.8 thousand m^2 - 4.8 thousand m^2), respectively. The corresponding estimate for stratum 4 was very imprecise at 19 thousand m^2 (1.6 thousand m^2 - 46.5 thousand m^2). No surface oil was found in strata 5 and 6. The grand sum for Prince William Sound is 41 thousand m^2 (20.7 thousand m^2 - 70.5 thousand m^2) (Table 5).

Estimates of the total subsurface oil weight (metric tons, or 1000s kg) per stratum and measures of their precision have been computed for Adaptive, stratified random sampling (SRS), and Observed modes, either from analytical formulas or by bootstrap resampling (Table 6). Observed estimates of oil weight, when extrapolated to stratum totals, provide minimal, biased values. Considering Adaptive estimates and bootstrap computations as most reliable, total weights of subsurface oil for strata 1 through 6 were estimated at 8 m.t. (2.7 m.t. - 13.8 m.t.), 18 m.t. (6.3 m.t.- 34.7 m.t.), 4 m.t. (0.3 m.t. - 10.9 m.t.), 17 m.t. (0.5 m.t. - 42.9 m.t.), 4 m.t. (0 m.t. - 11.3 m.t.), and 0 m.t. (none observed), respectively. The grand sum of subsurface oil weight for Prince William Sound was 50 m.t. (24.4 m.t. - 82.6 m.t.) (Table 7).

Although the added effort for adaptive pits reduced uncertainty for the oiled areas of sampled subsegments, the variation among subsegments within strata greatly increased the variances for stratum totals. For example, the variance equation for stratum total oiled area in 100 m subsegments (see Equation 2) comprises two terms, the first representing variation among subsegments, and the second, estimation error within subsegments. In applying Equation 2 to Adaptive estimates of strata 2 and 4, the contribution to variance of stratum total oiled area from variation among subsegments represented more than 99% of total variance. Because the within segment variation was negligible as compared to between segment variation, the adaptive sampling would better have been implemented by allocating sampling effort among strata from oil found in a preliminary exploration, instead of delimiting the oil patches within an oiled beach segment. For example, stratum 4 contained roughly 40% of all estimated subsurface oil (Tables 2 and 3) whereas only 3% of the segments were sampled in that stratum(Table 1). By implementing an adaptive program on the strata with additional beach segments sampled from strata where more oil was found, precision in estimation of the amount of Exxon Valdez spill oil remaining in Prince William Sound would have been enhanced. Unfortunately, the permits required before sampling began prevented such adaption to discovery.

The bootstrap procedure, although computer intensive, provided more realistic measures of precision evident from the asymmetric 95% confidence intervals from skewed distributions of the bootstrap estimates. The minimal oil estimate should never be less than zero, as occurred with the analytical mode, nor less than the amount of oil actually found. The bootstrap calculations were more consistent with this principle than the analytical calculations. None of the lower bounds from the bootstrap were negative, and only when the oil was very scarce did bootstrap resampling fail to produce a lower bound at least as great as the found oil. In implementing the bootstrap resampling, the negligible estimation error within 100 m subsegments, but not <100 m subsegments, was omitted from the computations. If sampling effort were better distributed among oiling strata in future surveys, the relative importance of estimation error within subsegments could increase and necessitate its inclusion in bootstrap calculations.

Literature Cited

Booth, J. G., R. W. Butler, and P. Hall. 1994. Bootstrap methods for finite populations. J. Amer. Statistical Association 89: 1282-1289.

Brodersen, C., Short, J., Holland, L., Carls, M., Pella, J., Larsen, M., Rice, S. Evaluation of oil removal from beaches 8 years after the Exxon Valdez oil spill. Proc. 22nd Arctic & Marine Oilspill Program Technical Seminar, Environment Canada, Ottawa. 1999. pp. 325-336.

Cochran, W. G. 1963. Sampling techniques. Second Edition. John Wiley & Sons, Inc., New York, 413p.

Efron, B. 1982. The jackknife, the bootstrap, and other resampling methods, Volume 38, CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics, Philadelphia..

Finkelstein, K., Gundlach, E. R. Method for estimating spilled oil quantity on the shoreline. Environ. Sci. Technol. 15, 545-549 (1981).

Gibeaut, J.C., Piper, E., Munson, D., Matthews, J., Profita, M., Crosby, C. S. 1993 shoreline oiling assessment of the *Exxon Valdez* oil spill. *Exxon Valdez* Oil Spill Restoration Project Final Report (Restoration Project 93038), Alaska Department of Environmental Conservation, Juneau, Alaska (1998).

Goodman, L. A. 1960. On the exact variance of products. J. Amer. Statist. Assoc. 55: 708-713. (Corr. 57: 917).

Gross, S. 1980. Median estimation in sample surveys. Proceedings of the Section on Survey Research Methods. American Statistical Association, pp. 181-184.

Koons, C. B., and H. O. Jahns. 1992. The Fate of Oil from the Exxon Valdez - A Perspective. Journal Marine Technology Society. 61-69.

Neff, J. M., Owens, E. H., Stoker, S. W. Shoreline oiling conditions in Prince William Sound following the *Exxon Valdez* oil spill. Pp. 312-346 in Wells, P. G., Butler, J. N., Hughs, J. S., eds., *Exxon Valdez Oil Spill: Fate and Effects in Alaskan Waters*. American Society for Testing and Materials Publication 1219, ASTM, Philadelphia, PA. (1996).

Owens, E. H. SCAT – A ten-year review. *Proc. 22nd Arctic & Marine Oilspill Program Technical Seminar*, Environment Canada, Ottawa pp. 337-360.

Short, J. W., Lindeberg, M. R., Harris, P.M., Maselko, J., Rice, S. D. Vertical Oil Distribution within the Intertidal Zone 12 Years after the Exxon Valdez Oil Spill in Prince William

Sound Alaska. *Proc. 25th Arctic & Marine Oilspill Program Technical Seminar*, Environment Canada, Ottawa pp. 57-72.

Thompson, S. K. 1992. Sampling. Wiley and Sons, New York, 343p.

Appendix I

The estimation equation is $\hat{T}_1 = N_1 \left(\frac{\sum_{i=1}^{n_1} \hat{y}_{1i}}{n_1} \right)$, and its variance is given by the general formula (see Equation 10 on page 134 of Thompson, 1992),

$$Var(\hat{T}_1) = \text{var}\left[E(\hat{T}_1 | s_1)\right] + E\left[\text{var}(\hat{T}_1 | s_1)\right].$$

Here s_1 denotes one of the $\binom{N_1}{n_1}$ possible selections of n_1 100 m beach subsegments from the N_1 available. The conditional expectation and variance of $\hat{T}_1$, given a particular selection of 100 m beach subsegments, s_1, are taken over the possible samples of quadrats obtained by adaptive sampling in s_1. The unconditional variance and expectation are taken over all possible selections of s_1. For any 100 m beach subsegment included in sampling (say the ith subsegment), the expected value of $\hat{y}_i$ is y_i because the adaptive stratified sampling is unbiased. Therefore,

$$E\left[(\hat{T}_1 | s_1)\right] = N_1 \left(\frac{\sum_{i=1}^{n_1} E(y_i | s_1)}{n_1} \right) = N_1 \left(\frac{\sum_{i=1}^{n_1} y_i}{n_1} \right) = N_1 \bar{y}_1 .$$

Because of simple random sampling of the 100 m subsegments, the corresponding unconditional variance is

$$\text{var}\left[E(\hat{T}_1 | s_1)\right] = \text{var}\left[N_1 \bar{y}_1 | s_1\right] = N_1^2 \text{ var}(\bar{y}_1),$$

where σ_y^2 is the population variance of the unobservable y_i, $i=1,\ldots,N_1$.

The conditional variance of $\hat{T}_1$, given the selection of 100 m beach subsegments s_1, depends on the estimation variances for the oiled areas of the individual beach subsegments. Let the estimation variance for the ith 100 m beach subsegment be denoted as

$$\sigma_{\hat{y}_i}^2 = E(\hat{y}_i - y_i)^2 .$$

Because the estimation errors are independent among beach subsegments,

$$\text{var}(\hat{T}_1 | s_1) = N_1^2 \text{ var} \left(\frac{\sum_{i=1}^{n_1} \hat{y}_i}{n_1} \right) = \left(\frac{N_1}{n_1} \right)^2 \sum_{i=1}^{n_1} \text{var}(\hat{y}_i).$$

The unconditional expected value of this conditional variance over all possible selections of 100 m beach subsegments is obtained by setting z_i to be an indicator variable for presence of the ith beach subsegment in the random sample (z_i=1 if present, and z_i=0 otherwise) and rewriting the preceding conditional variance as

$$\text{var}(\hat{T}_1|s_1) = \left(\frac{N_1}{n_1}\right)^2 \sum_{i=1}^{N_1} z_i \text{ var}(\hat{y}_i).$$

Because the expected value of z_i=n_1/N_1, the expected value of the conditional variance is

$$E(\text{var}(\hat{T}_1|s_1)) = \sum_{i=1}^{N} \left(\frac{N_1}{n_1}\right)^2 E(z_i)\text{ var}(\hat{y}_i) = \left(\frac{N_1}{n_1}\right)\sum_{i=1}^{N_1} \text{var}(\hat{y}_i).$$

Appendix II

Notice that the variance, $V(\hat{W}_1)=\sum_{j=1}^{M}V(\overline{A}_j\overline{w}_j)+2\sum_j\sum_{k<j}Cov(\overline{A}_j\overline{w}_j,\overline{A}_k\overline{w}_k)$. We assume reasonably that $\overline{A}_j$ and $\overline{w}_j$ are independent, so that (Goodman 1960)

$$V(\overline{A}_j\overline{w}_j)=\mu_{wj}^2V(\overline{A}_j)+A_j^2V(\overline{w}_j)+V(\overline{A}_j)V(\overline{w}_j).$$

Also, for different visual categories, the average weights, $\overline{w}_j$ and $\overline{w}_k$, are independent and so it can be shown by using the definition for covariance that

$$Cov(\overline{A}_j\overline{w}_j,\overline{A}_k\overline{w}_k)=\mu_{wj}\mu_{wk}Cov(\overline{A}_j,\overline{A}_k).$$

The variance of the visual category means, $\overline{w}_j$, is $V(\overline{w}_j)=\sigma_{wj}^2/n_{wj}$. Therefore, the variance of $\hat{W}_1$ is obtained by summing terms, *viz*,

$$V(\hat{W}_1)=\sum_{i=1}^{M}\left\{A_i^2\frac{\sigma_{wi}^2}{n_{wi}}+\mu_{wi}^2V(\overline{A}_i)+V(\overline{A}_i)\frac{\sigma_{wi}^2}{n_{wi}}\right\}+2\sum_{i=1}^{M}\sum_{j<i}\mu_{wi}\mu_{wj}Cov(\overline{A}_i,\overline{A}_j)$$

Appendix III

The estimate of visual areas in the kth subsegment is $\mathbf{\hat{a}_k} = \hat{y}_k \cdot \begin{pmatrix} \hat{p}_{k1} \\ \vdots \\ \hat{p}_{kM} \end{pmatrix} = \hat{y}_k \cdot \mathbf{\hat{p}_k}$, where $\hat{y}_k$ is the estimated oiled area in beach subsegment k (a scalar) and $\mathbf{\hat{p}_k}$ is the estimated visual area composition of subsegment k (a vector). The two terms of $\mathbf{\hat{a}_k}$ are reasonably considered independent. The jth diagonal element of the variance-covariance matrix of estimation errors for the kth beach subsegment is defined as

$$v_{\hat{a}_k - a_k, j, j} = E\left[(\hat{y}_k \hat{p}_{kj})^2\right] - \left[E(\hat{y}_k \hat{p}_{kj})\right]^2.$$

A new expression is obtained by adding and subtracting the term, $\left[E(\hat{y}_k)\right]^2 E(\hat{p}_{kj}^2)$, and simplifying the result under the assumption that $\hat{y}_k$ and $\hat{p}_{kj}$ are unbiased as well as independent, *viz.*,

$$v_{\hat{a}_k - a_k, j, j} = E(\hat{y}_k^2)E(\hat{p}_{kj}^2) - \left[E(\hat{y}_k)\right]^2 E(\hat{p}_{kj}^2) + \left[E(\hat{y}_k)\right]^2 E(\hat{p}_{kj}^2) - \left[E(\hat{y}_k)\right]^2 \left[E(\hat{p}_{kj})\right]^2$$

$$= E(\hat{p}_{kj}^2)\sigma_{\hat{y}_k}^2 + \left[E(\hat{y}_k)\right]^2 \frac{p_{kj}(1-p_{kj})}{h_k} = \left(y_k^2 + \sigma_{\hat{y}_k}^2\right) \cdot \frac{p_{kj}(1-p_{kj})}{h_k} + \sigma_{\hat{y}_k}^2 p_{kj}^2$$

The off-diagonal element at the ith row and jth column is obtained by similar expansion of the definition,

$$v_{\hat{a}_k - a_k, i, j} = E(\hat{y}_k \hat{p}_{ki} \hat{y}_k \hat{p}_{kj}) - E(\hat{y}_k \hat{p}_{ki})E(\hat{y}_k \hat{p}_{kj}) = E(\hat{y}_k^2)E(\hat{p}_{ki}\hat{p}_{kj})$$

$$- E(\hat{y}_k^2)E(\hat{p}_{ki})E(\hat{p}_{kj}) + E(\hat{y}_k^2)E(\hat{p}_{ki})E(\hat{p}_{kj}) - \left[E(\hat{y}_k)\right]^2 E(\hat{p}_{ki})E(\hat{p}_{kj})$$

$$= -\left(y_k^2 + \sigma_{\hat{y}_k}^2\right)\frac{p_{ki}p_{kj}}{h_k} + \sigma_{\hat{y}_k}^2 p_{ki} p_{kj}$$

List of Figures

List of Tables

Table 1. Strata sizes in numbers and total lengths of segments, sample sizes, and the sampling fractions.

Stratum (h)	Segments (N_h)	Sampled segments (n_h)	Sampling fraction (%)	Total length (m)	Sampled length (m)	Sampling fraction of length (%)
1. 1990-93, heavy, <100 m	133	23	17	5,935	1,576	27
2. 1990-93, heavy, 100 m	184	37	20	18,458	3,708	20
3. 1990-93, med, < 100 m	310	16	5	12,827	978	8
4. 1990-93, med, 100 m	362	10	3	36,303	1,000	3
5. 1989, heavy, < 100 m	221	5	2	11,294	253	2
6. 1989, heavy, 100 m	317	5	2	31,775	500	2
Totals	1,527	96	6	116,592	8,015	7

Table 2. Subsurface area (in thousands m^2) estimation results per stratum comparing analytical and bootstrap calculations. Note that no oil was found in Strata 6, 1989 heavy oil 100 m beach segments. The column, "estimate", refers to point estimate if calculation mode is analytical, or mean estimate, if bootstrap.

Stratum (h)	Computation method	Estimate	Lower 95% CI	Upper 95% CI	Standard Error	Coefficient of Variation
1. 1990-93, heavy, <100 m	Analytical Adaptive	13.4	3.0	23.9	5.35	0.40
	Bootstrap Adaptive	13.3	4.4	24.6	5.18	0.39
	Analytical SRS	12.8	3.7	21.9	4.64	0.36
	Bootstrap SRS	12.9	4.9	23.2	4.63	0.36
	Analytical Observed	10.4	0.4	20.3	5.09	0.49
	Bootstrap Observed	10.6	2.5	21.3	4.95	0.47
2. 1990-93, heavy, 100 m	Analytical Adaptive	21.8	9.2	34.5	6.46	0.30
	Bootstrap Adaptive	21.6	10.3	36.0	6.61	0.31
	Analytical SRS	22.3	10.4	34.2	6.05	0.27
	Bootstrap SRS	22.6	11.4	36.7	6.48	0.29
	Analytical Observed	15.5	3.8	27.2	5.97	0.38
	Bootstrap Observed	15.6	4.8	29.3	6.31	0.41
3. 1990-93, med, < 100 m	Analytical Adaptive	5.8	-1.6	13.2	3.78	0.65
	Bootstrap Adaptive	5.8	0.9	13.8	3.73	0.64

	Analytical SRS	6.9	-1.5	15.4	4.31	0.62
	Bootstrap SRS	7.1	0.9	16.6	4.20	0.59
	Analytical Observed	1.8	-0.6	4.1	1.20	0.67
	Bootstrap Observed	1.7	0.07	4.2	1.14	0.65
4. 1990-93, med, 100 m	Analytical Adaptive	26.6	-7.9	61.1	17.6	0.66
	Bootstrap Adaptive	27.3	2.4	63.9	17.1	0.63
	Analytical SRS	29.8	-10.9	70.4	20.7	0.70
	Bootstrap SRS	29.0	2.4	71.4	20.4	0.70
	Analytical Observed	18.3	-16.5	53.1	17.8	0.97
	Bootstrap Observed	18.0	0.05	54.4	16.8	0.94
5. 1989, heavy, < 100 m	Analytical Adaptive	3.2	-3.1	9.5	3.22	1.00
	Bootstrap Adaptive	3.2	0.0	9.7	2.91	0.91
	Analytical SRS	6.6	-6.4	19.7	6.65	1.00
	Bootstrap SRS	6.4	0.0	19.9	6.25	0.98
	Analytical Observed	2.6	-2.5	7.7	0.870	1.00
	Bootstrap Observed	3.2	0.0	9.4	0.955	0.89

Table 3. Total Prince William Sound subsurface oil area (in thousands m^2) results comparing analytical and bootstrap approaches as well as adaptive, stratified random sample (SRS) and calculations based on observed oil area on selected beach segments. The column, "estimate", refers to point estimate if calculation mode is analytical, or mean estimate, if bootstrap.

Computation method	Estimate	Lower 95% CI	Upper 95% CI	Standard Error	Coefficient of Variation
Analytical Adaptive	71.0	31.5	110.4	20.1	0.28
Bootstrap Adaptive	71.2	37.7	113.2	19.7	0.28
Analytical SRS	78.4	32.4	124.4	23.5	0.30
Bootstrap SRS	78.0	40.6	127.3	22.8	0.29
Analytical Observed	48.6	10.2	87.1	19.6	0.40
Bootstrap Observed	49.1	19.9	92.7	18.7	0.38

Table 4. Surface area (in thousands m^2) estimations per strata comparing analytical and bootstrap calculations. Note that no surface oil was found in Strata 5 and 6 (1989 heavy oil <100m and 100m beach segments). The column, "estimate", refers to point estimate if calculation mode is analytical, or mean estimate, if bootstrap.

Stratum (h)	Computation method	Estimate	Lower 95% CI	Upper 95% CI	Standard Error	Coefficient of Variation
1. 1990-93, heavy, <100 m	Analytical SRS	8.8	3.9	13.7	2.49	0.28
	Bootstrap SRS	7.6	3.3	12.7	2.44	0.32
2. 1990-93, heavy, 100 m	Analytical SRS	12.0	5.9	18.0	3.10	0.26
	Bootstrap SRS	12.1	6.3	19.0	3.29	0.27
3. 1990-93, med, < 100 m	Analytical SRS	2.7	0.6	4.7	1.05	0.39
	Bootstrap SRS	2.6	0.8	4.8	1.02	0.39
4. 1990-93, med, 100 m	Analytical SRS	19.4	-6.1	45.0	13.0	0.67
	Bootstrap SRS	19.0	1.6	46.5	12.5	0.66

Table 5. Total Prince William Sound surface oil area (in thousands m^2) results comparing analytical and bootstrap calculations. The computations were made using stratified random sample as no adaptive sampling was done for surface area calculations. The column, "estimate", refers to point estimate if calculation mode is analytical, or mean estimate, if bootstrap.

Computation method	Estimate	Lower 95% CI	Upper 95% CI	Standard Error	Coefficient of Variation
Analytical SRS	42.9	16.1	69.7	13.7	0.32
Bootstrap SRS	41.3	20.7	70.5	13.2	0.32

Table 6. Subsurface oil weight (in metric tons) per strata comparing analytical and bootstrap calculations as well as adaptive and observed oil area. The column, "estimate", refers to point estimate if calculation mode is analytical, or mean estimate, if bootstrap.

Stratum (h)	Computation method	Estimate	Lower 95% CI	Upper 95% CI	Standard Error	Coefficient of Variation
1. 1990-93, heavy, <100 m	Analytical Adaptive	7.7	1.9	13.5	2.95	0.38
	Bootstrap Adaptive	7.6	2.7	13.8	2.88	0.38
	Bootstrap SRS	7.5	3.0	13.8	2.71	0.36
	Bootstrap Observed	5.9	1.6	12.2	2.71	0.46
2. 1990-93, heavy, 100 m	Analytical Adaptive	18.3	4.1	32.5	7.25	0.40
	Bootstrap Adaptive	18.0	6.3	34.7	7.49	0.42
	Bootstrap SRS	18.7	6.7	36.6	7.59	0.41
	Bootstrap Observed	14.5	3.6	31.0	7.00	0.48
3. 1990-93, med, < 100 m	Analytical Adaptive	4.0	-1.7	9.6	2.88	0.73
	Bootstrap Adaptive	4.0	0.3	10.9	2.92	0.74
	Bootstrap SRS	4.9	0.3	12.4	3.27	0.66
	Bootstrap Observed	1.3	0.02	3.3	0.88	0.70
4. 1990-93, med, 100 m	Analytical Adaptive	16.0	-6.6	38.6	11.5	0.72
	Bootstrap Adaptive	16.5	0.5	42.9	11.4	0.69
	Bootstrap SRS	17.5	0.5	48.8	13.6	0.78

	Bootstrap Observed	11.5	0.01	37.0	11.0	0.96
5. 1989, heavy, < 100 m	Analytical Adaptive	3.5	-3.4	10.4	3.53	1.00
	Bootstrap Adaptive	3.5	0.0	11.3	3.30	0.94
	Bootstrap SRS	7.0	0.0	24.1	7.03	1.00
	Bootstrap Observed	3.5	0.0	10.7	3.21	0.91

Table 7. Total Prince William Sound subsurface oil weight (in metric tons) comparing analytical and bootstrap calculations as well as adaptive and observed oil area. The column, "estimate", refers to point estimate if calculation mode is analytical, or mean estimate, if bootstrap.

Computation method	Estimate	Lower 95% CI	Upper 95% CI	Standard Error	Coefficient of Variation
Analytical Adaptive	49.5	22.4	76.7	13.8	0.28
Bootstrap Adaptive	49.6	24.4	82.6	15.2	0.31
Bootstrap SRS	55.6	26.1	94.4	17.9	0.32
Bootstrap Observed	36.7	14.5	66.4	14.0	0.38

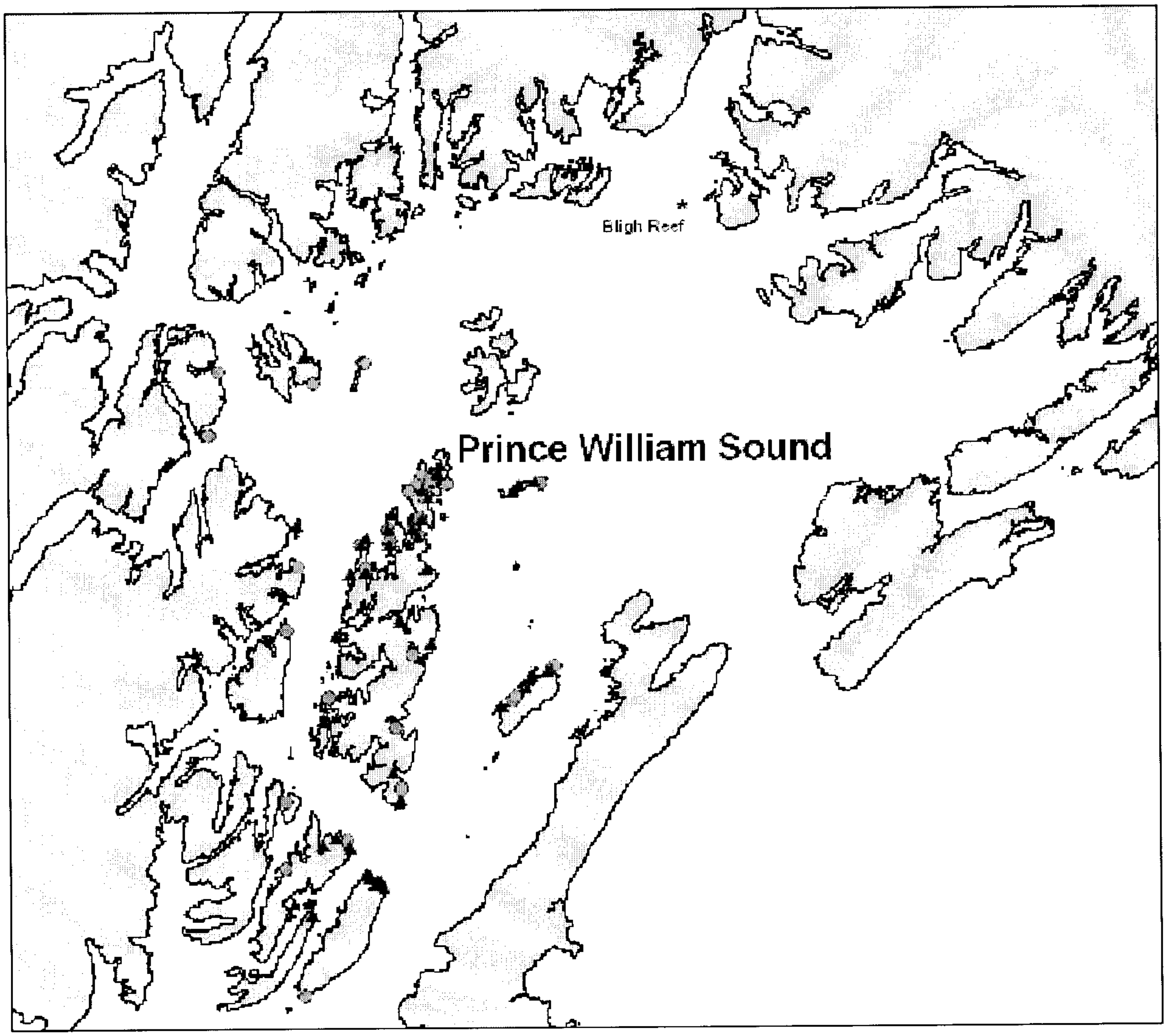


Figure 1. Prince William Sound with Bligh Reef (*), the grounding site of the Exxon Valdez, and the beaches affected by the spill. Symbols indicate stations where oil was detected (triangles) or not (circles) in our randomly placed quadrats.

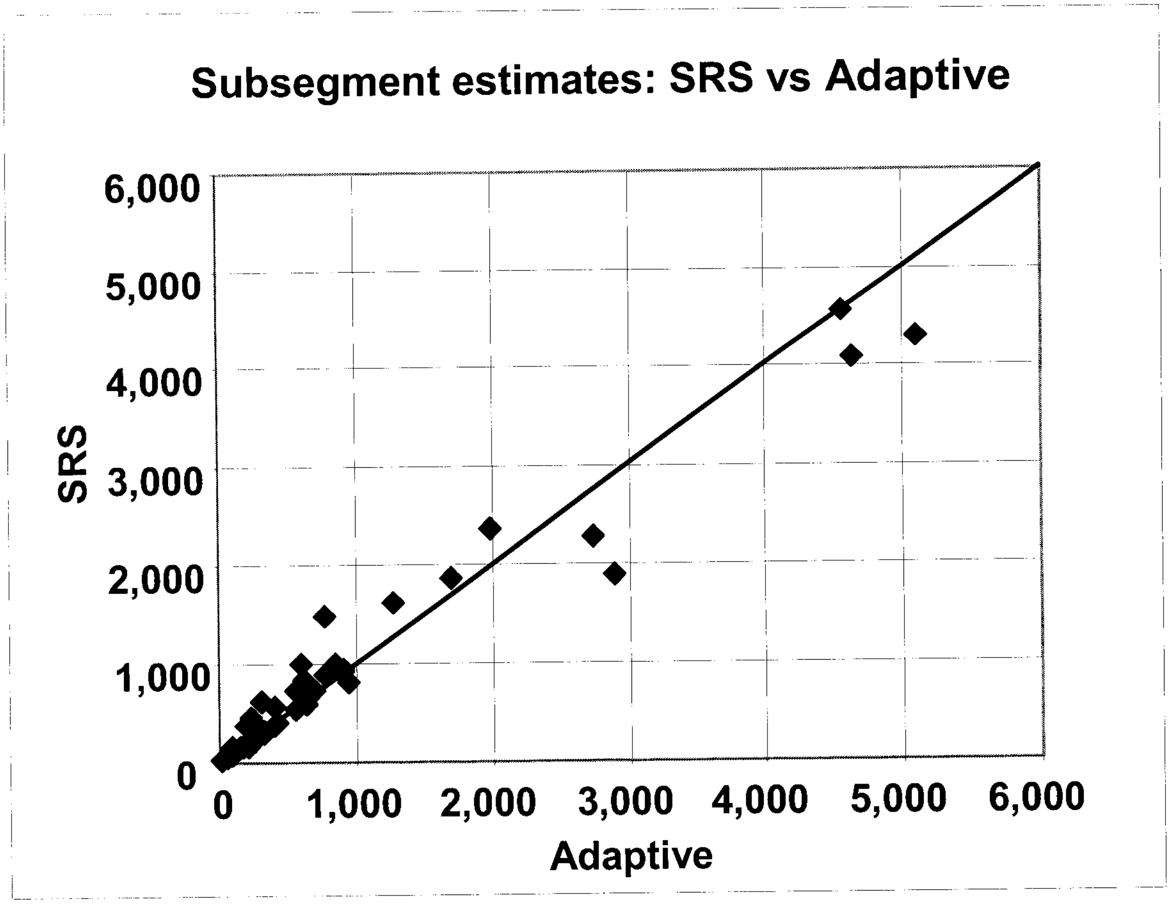


Figure 2. Estimated subsurface oiled areas (in quadrats) in sampled subsegments by stratified random sampling (SRS) and Adaptive modes, and the line of equality between modes.

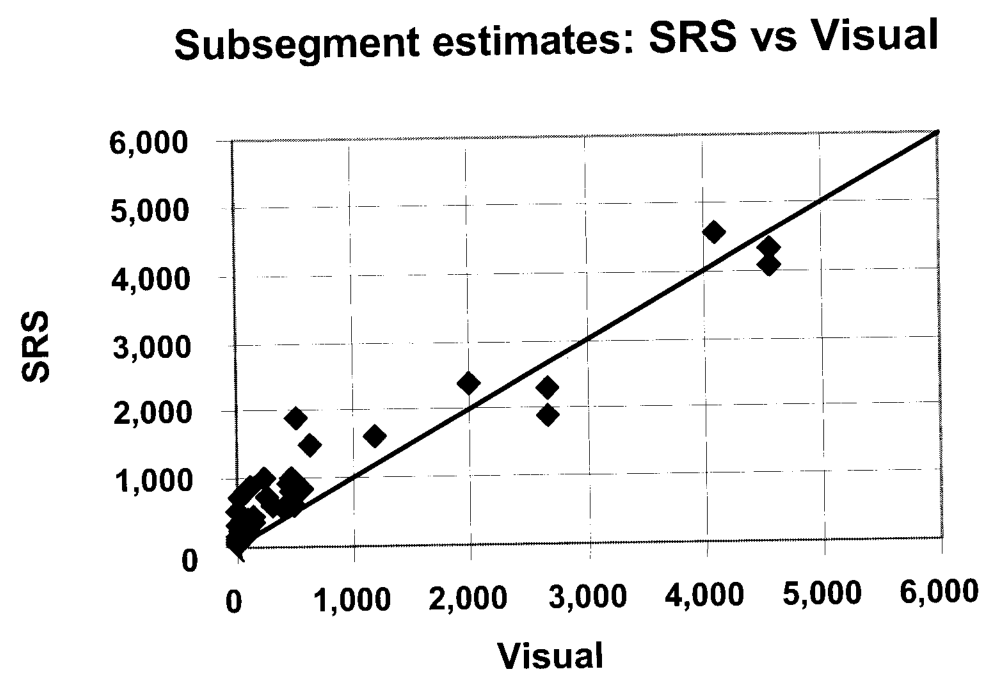


Figure 3. Estimated subsurface oiled areas (in quadrats) in sampled subsegments by stratified random sampling (SRS) compared to the numbers of sampled quadrats with oil present (Visual). Line of equality separates known underestimates (below line) from feasible estimates (above line).

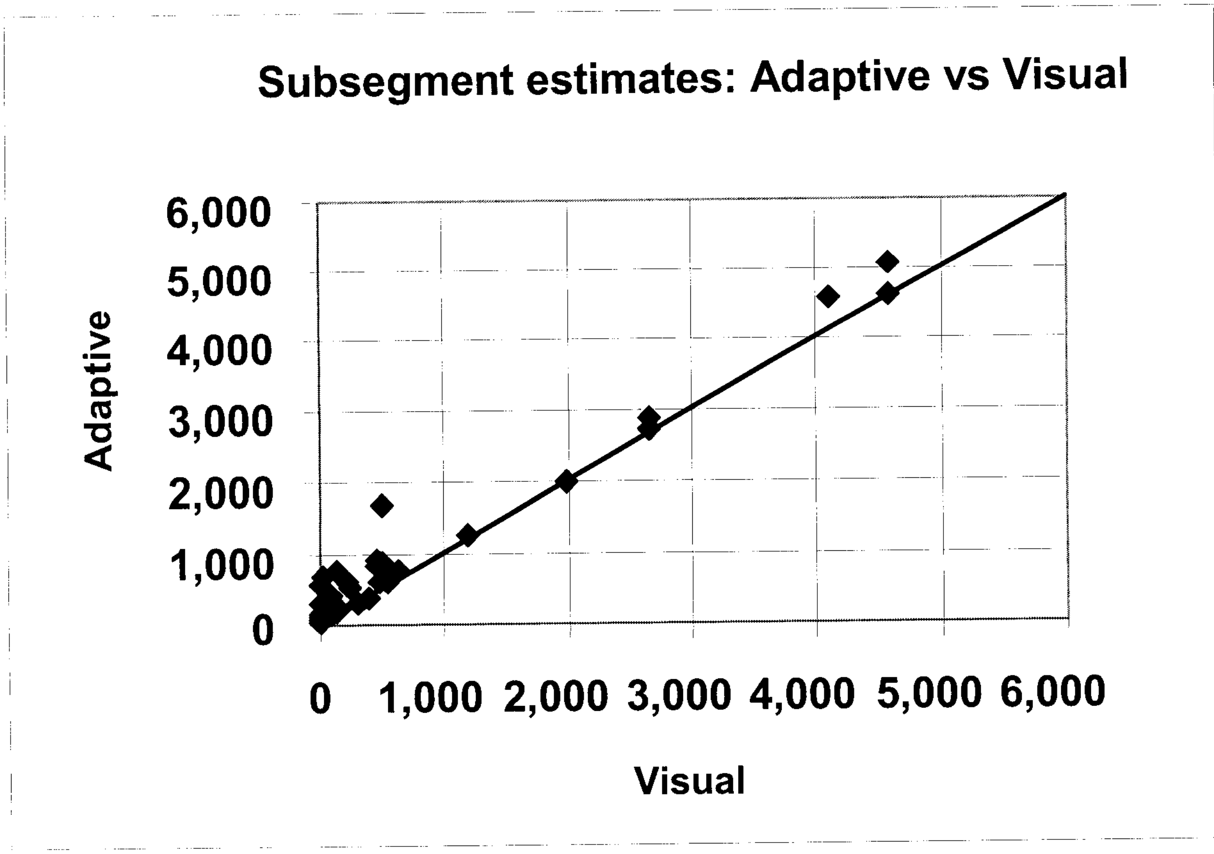


Figure 4. Estimated subsurface oiled areas (in quadrats) in sampled subsegments by Adaptive sampling compared to the numbers of sampled quadrats with oil present (Visual), and line of equality.

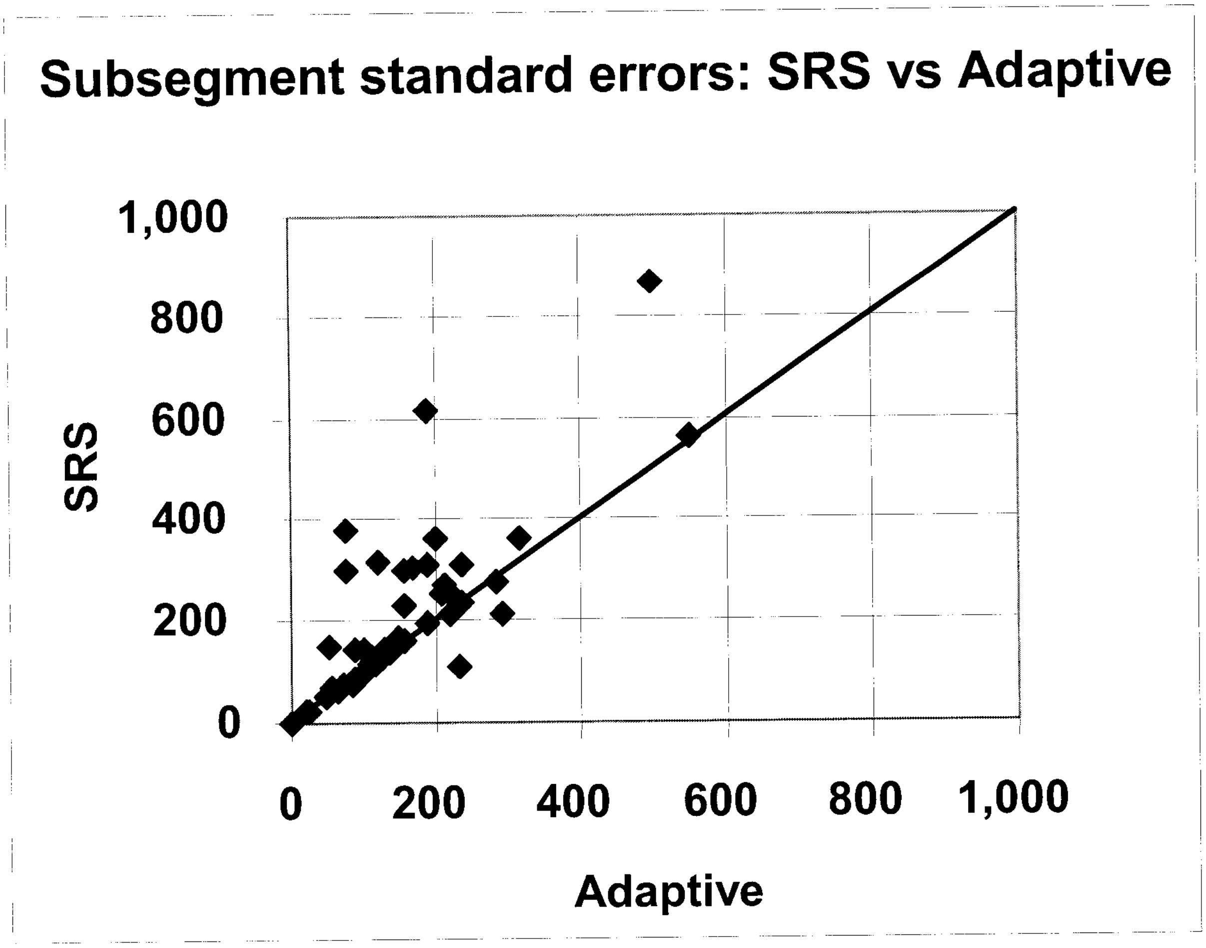


Figure 5. Estimated standard errors of subsurface oiled area estimates for sampled subsegments, by stratified random sampling (SRS) and Adaptive modes, and line of equality.

Dispersion of Crude Oil from the *Exxon Valdez* from Subtidal Sediments of Prince William Sound, Alaska, 12 Years after the Event

Jeffrey W. Short* and Stanley D. Rice

Auke Bay Laboratory, Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA, 11305 Glacier Highway, Juneau, Alaska 99801-8626

*Corresponding author telephone: 907-789-6065; fax: 907-789-6094; e-mail: Jeff.Short@noaa.gov

Abstract

We sampled inter- and subtidal sediments from one reference and four oiled locations in Prince William Sound, Alaska, during summer 2001, to evaluate the persistence of remnant oil from the 1989 *Exxon Valdez* oil spill. At each location, sediment samples were collected from seven depths along transects beginning at 0 m (Mean Lower Low Water) and ending at 100 m, at stations that had been repeatedly sampled during surveys from 1989 to 1991. The collected sediments were analyzed for a suite of aliphatic and polycyclic aromatic hydrocarbons (PAH) using methods directly comparable with those used during the 1989 - 1991 surveys. Oil concentrations found in sub-tidal surface sediments were very low, lower than the concentrations of remnant *Exxon Valdez* oil found in 1989-91 or in the subsurface pits found in 2001 in the intertidal zone. Subtidal oil concentrations from the oiled and reference locations were similar, and showed PAH concentrations that generally increased with depth from total PAH concentrations near detection limits of a few ng/g at 0 m to 776 ng/g at 100 m. Source of PAH in subtidal sediments was not *Exxon Valdez* oil; trend of concentrations increasing with depth was not consistent with the findings in 1989-91, and most importantly, the pattern of PAH in the samples was not consistent with weathered *Exxon Valdez* oil. Most of these PAH were from the regional background of coal and organic-rich shales that erode from the terrestrial Kulthieth Formation east of the Copper River and transported into PWS by the Alaska Coastal Current. Small contributions from anthropogenic sources were likely present at three stations, and trace contributions from *Exxon Valdez* oil may be present at two, but the evidence for this is weak. Because the four locations we sampled in western PWS during 2001 were among the most heavily contaminated by the *Exxon Valdez* oil spill, we believe contributions from remnant *Exxon Valdez* oil to the subtidal sediments are now negligible in comparison with contributions from the regional background.

Introduction

About 40% of the oil spilled from the T/V *Exxon Valdez* oiled shorelines of Prince William Sound (PWS; Wolfe et al. 1994), creating a reservoir of sequestered oil that could be slowly re-introduced into adjacent marine waters. As a result of various beach-cleaning efforts during 1989 and 1990, or of exposure to high-energy waves during winter storms, some of this oil subsequently migrated to subtidal sediments, especially adjacent to heavily oiled beaches (O'Clair et al. 1996, Short et al. 1996a). During the first few years following the spill, migration of this oil to the subtidal, as coatings on fine-grained sediments, was clearly demonstrated in sediment trap studies (Short et al. 1996a), as was contamination of nearshore sediments, especially at seawater depths shallower than 20 m (O'Clair et al. 1996). Traces of oil were detected at depths exceeding 100 m in some instances (Carlson & Kvenvolden, 1996). Studies of subtidal geomorphology indicated that tidal and storm-surge currents were often strong enough to disperse fine-grained, oil-coated sediments to progressively deeper depths of PWS, diluting them to concentrations that were likely undetectable in the process (Short et al. 1996a). The most important exception was oiled sediments initially deposited in subtidal basins that could act as a catchment for the sediments, as at Northwest Bay (Short et al., 1996a). But in any case, the long-term persistence of oil in subtidal sediments of PWS has not been reliably monitored since 1991.

Reports of oil persisting on PWS beaches (Brodersen et al., 1999; Carls et al., 2001) raised the possibility of continued introductions of low concentrations of fine-grained, oiled sediments to the adjacent subtidal. As part of a quantitative study to estimate the amount of oil remaining in PWS in 2001 (Short et al., 2004), the study we report here examined hydrocarbon concentrations in sediments adjacent to five beaches to determine whether *Exxon Valdez* oil was still evident in them. The five selected beaches were chosen because their adjacent subtidal sediments had been sampled extensively during surveys from 1989 through 1991, along transects from the intertidal to 100 m depth (O'Clair et al., 1996). Four of these beaches were among the hardest hit beaches during the spill. The sediment samples collected from re-occupation of these transects were analyzed by the same methods as were used for the samples collected previously in 1989 to 1991, permitting direct comparison of results. Comparison of these results allows evaluation of oil inputs, persistence and dispersion during the intervening ten years from 1991 to 2001.

Methods

Sampling Stations

The five sampling locations selected for this study include one un-oiled on the east side of PWS at Olsen Bay, and four locations that were heavily oiled initially at Herring Bay, Northwest Bay, Sleepy Bay and Snug Harbor. These locations correspond with location numbers 32, 20, 31, 38, and 41, respectively, in O'Clair et al. (1996). At each location, samples were collected along transects at stations where seawater depths were 0 m (mean lower low water; MLLW), 3 m, 6 m, 10 m, 20 m, 40 m and 100 m below MLLW. These transects were located based on co-ordinates and maps used for the 1989 - 1991 O'Clair et al. study (1996). We estimate our re-occupation of these stations is within about 10 m of the original stations, based on uncertainties implied by the mapping precision of the 1996 O'Clair et al. study.

Sediment Sample Collection
The sediment samples were collected with a Ponar sampler during 18 - 20 June, 2001, aboard the M/V Auklet. The sampler had a maximum sampling depth of 4 cm. The sampler was partially opened after retrieval to allow water drainage, and the collected sediments were examined for evidence of an un-disturbed surficial layer, indicated by intact siphon holes or an intact algal layer. If an undisturbed layer was present, about 100 g wet mass of the topmost 1 cm. of sediment from the un-disturbed layer was removed with a dichloromethane-rinsed steel spoon and deposited in a pre-cleaned glass jar equipped with a PFTE-lined lid; otherwise the sample was rejected and another sample collected. Collected samples were immediately frozen aboard ship, and were transported frozen to the Auke Bay Laboratory where they were stored at - 20 C until hydrocarbon analysis.

Hydrocarbon Analysis
Sediment samples (2 g – 15 g dry wt.) were extracted with dichloromethane at 100 °C and 2000 psi for 10 min in a Dionex ASE 200 accelerated solvent extractor. The dichloromethane solutions were exchanged with hexane over steam, and separated into aliphatic and aromatic fractions by column chromatography (10 g 2%-deactivated alumina over 20 g 5%-deactivated silica gel). Aliphatics eluting with 50 mL pentane were analyzed by GC-FID following concentration to ~ 1 mL hexane over steam. Polycyclic aromatic hydrocarbons (PAH) eluting with 50 mL 1:1 (v/v) pentane:dichloromethane were exchanged into ~ 1 mL hexane over steam, and further purified by size-exclusion high-performance liquid chromatography. Purified PAH were measured by GCMS operated in the selected ion monitoring mode.

The aliphatic hydrocarbons analyzed include the *n*-alkanes from *n*-nonane (nC_9) through *n*-tetratriacontane (nC_{34}), pristane (2,6,10,14-tetramethylhexadecane; abbreviated as Pr) and phytane (2,6,10,14-tetramethylheptadecane; Ph). Aliphatic hydrocarbon concentrations were determined using an internal-standard method employing a five-point calibration curve for each target analyte, with the perdeuterated surrogate standards treated as the internal standards.

The aromatic hydrocarbons analyzed included both parent homolog and alkyl-substituted naphthalenes, fluorenes, dibenzothiophenes, phenanthrenes, and chrysenes (abbreviated as NX, FX, DX, PX and CX, where X = number of alkyl-substituted carbon atoms ranging from 0 to 3 or 4), along with biphenyl (bip), acenaphthene (ace), acenaphthylene (acn), anthracene (ant), fluoranthene (flu), pyrene (pyr), methyl-fluoranthene/pyrene (FP1), benzo-a-anthracene (baa), benzo-b-fluoranthene (bbf), benzo-k-fluoranthene (bkf), benzo-e-pyrene (bep), benzo-a-pyrene (bap), perylene (per), indeno-1,2,3-*c,d*-pyrene (icp), dibenzo-*a,h*-anthracene (dba), and benzo-*g,h,i*-perylene (bgp). The concentrations of N1 and the unsubstituted PAHs were determined in a manner similar to that used for the aliphatic hydrocarbons. Calibration curves for the most similar PAHs were used to calculate concentrations of the other alkyl-substituted PAHs: 2,6-dimethylnaphthalene for N2, 2,3,5-trimethylnaphthalene for N3 and N4, 1-methylphenanthrene for P1-P4 and otherwise the unsubstituted homolog.

Samples were extracted and analyzed in batches of 12 along with two reference samples, a method blank and a spiked method blank for quality control. Method detection limits (MDLs) varied inversely with sample aliquot size and are summarized in Short et al. (1996b). These MDLs are about 1 ng g^{-1} dry wt. sediment for the PAH and about 20 ng g^{-1} for the aliphatics. Sample precision is generally better than ± 20% based on the combined results of the reference samples, and accuracy is generally better than ± 15% based on comparison with National Institute of Standards and Technology standards used for the spiked blanks.

Hydrocarbon concentrations are reported on a dry weight basis. Dry weights of the sediment samples were determined by drying 1 g of wet sediment at 60 C for 24 h.

Data Analysis
Total polycyclic aromatic hydrocarbons (TPAH) is the sum of the PAH analyzed, excluding perylene, which is diagenetically produced from recent organic material. Relative concentrations of PAH are PAH concentrations divided by the TPAH, and are presented in figures 2 through 5.

Polycyclic Aromatic Hydrocarbon Sources
Patterns of relative PAH concentrations were compared with patterns characterizing two prospective hydrocarbon sources in this region: weathered *Exxon Valdez* oil, and organic-rich shales eroded from terrestrial occurrences east of PWS. The organic-rich shales are eroded by streams and glaciers between Yakutat and the Copper River, and are transported westward as fine-grained sediments entrained in the Alaska Coastal Current after fluvial transport to the Gulf of Alaska (Page et al. 1996). Distribution patterns of PAH found in the sediment samples were qualitatively compared with patterns that characterize weathered *Exxon Valdez* oil and sediments of the northern Gulf of Alaska. Quantitative comparisons, e.g. following Short and Heintz (1997) were precluded by the high frequency of PAH concentrations that were below MDLs.

Results

Aliphatic Hydrocarbons
Aliphatic hydrocarbon concentrations were low; most were below MDLs (method detection limits). Aliphatic hydrocarbons detected most often above MDLs include pristane and the n-alkanes containing an odd number of carbon atoms from nonadecane (nC_{19}) to nonacosane (nC_{29}), with heptacosane (nC_{27}) or nonacosane (nC_{29}) usually most abundant. Concentrations were highest in shallow (3 m - 20 m) sediments at Olsen Bay, where nC_{27} and nC_{29} ranged from 200 ng/g, and these two n-alkanes were occasionally found at 100 ng/g - 332 ng/g elsewhere.

At two intertidal (i.e. 0 m) stations (Northwest Bay and Sleepy Bay), the distribution of n-alkane concentrations was nearly uniform from hexadecane (nC_{16}) through triacontane (nC_{30}), at concentrations of about 35 ng/g and 100 ng/g respectively. Phytane was detected at the Northwest Bay, 0 m station at 33 ng/g, and was just below MDL at 142 ng/g at Sleepy Bay, 0 m. Phytane was not detected at the other stations and locations sampled. At Northwest Bay, pristane was present at concentrations ranging from 50 ng/g to 724 ng/g in sediments to 40 m depth, and lower concentrations (to 200 ng/g) were also found at Sleepy Bay to 20 m depth.

Polycyclic Aromatic Hydrocarbons concentrations

Concentrations of TPAH were low, consistently less than 800 ng/g, and tended to increase with increasing depth at all of the locations sampled (Figure 2). At 0 m to 6 m depths, total PAH concentrations are usually less than 100 ng/g ng, in contrast to the 100 m depth where total PAH concentrations range from 286 ng/g to 776 ng/g.

Polycyclic Aromatic Hydrocarbons composition patterns

None of the patterns of relative PAH concentrations found in the sediment samples we collected were consistent with weathered *Exxon Valdez* oil. Weathered *Exxon Valdez* oil typically has abundances of PAH homologues that increase with increasing alkyl substitution, and preferential losses of PAH that have fewer rings. These characteristics were not found in our samples.

The composition patterns of relative PAH concentrations at all depths at Olsen Bay are similar to those found from 20 m to 100 m depths at the other locations, and all these are similar to those typical of the northern Gulf of Alaska (compare Figures 3A - 3C). With three exceptions, the composition patterns that characterize stations shallower than 20 m at the four sampling locations in western PWS are consistent with the northern Gulf of Alaska pattern (Figure 3C), after allowing for concentrations below MDLs, which occurred frequently when TPAH concentrations were below 100 ng/g.

Composition of PAH in three shallow samples indicate additional sources; at Sleepy Bay, 0 m; Herring Bay, 3 m; and Northwest Bay, 10 m, where TPAH concentrations were 282 ng/g, 280 ng/g, and 491 ng/g.

At Herring Bay (3 m) and Northwest Bay (10 m), the PAH composition patterns evident in were dominated by the 4- and 5-ring PAH (Figure 4 A-B). The chrysenes are the most abundant PAH homologue series present, with homologue concentrations that decrease with increasing alkyl substitution. In contrast, the sediments at Sleepy Bay, 0 m are dominated by the 3- and 4-ring PAH, with the phenanthrenes/anthracenes and the fluoranthenes/pyrenes the most abundant homologues, and also with homologue concentrations that decrease with increasing alkyl substitution (Figure 5).

Discussion

Evidence of *Exxon Valdez* oil was generally lacking in the subtidal samples collected in 2001, in contrast to the samples collected in 1989-91. PAH concentrations were generally low in 2001; lower than intertidal samples where remnant EVO was found (Short et al. 2003), and lower in concentrations compared to the shallow samples where remnant EVO was found in 1989-91 by O'Clair et al (1996). The composition of PAH in samples collected in 2001 did not match qualitative to EVO, indicating other sources of PAH were more significant that EVO in 2001.

Concentrations of TPAH at the four sampling locations in western PWS (high impact oiling in 1989) were often higher during the 1989 - 1991 sampling than comparable samples collected during this study, and the PAH in the earlier samples were usually attributable to the presence of *Exxon Valdez* oil based on composition (O'Clair et al., 1996). Contributions from *Exxon Valdez* oil ranged from several hundred to several thousand ng/g, were typically highest in the shallow sediments, and declined with depth,. In contrast, the trend in 2001 is the opposite (increasing concentrations with depth, and is attributable to the natural background. Evidence for contributions from *Exxon Valdez* oil in the sediments collected during 2001 is scant, limited to perhaps some contribution to sediments at 0 m at Sleepy Bay and Northwest Bay, based on the phytane concentrations. Phytane is a persistent branched aliphatic that is characteristic of petroleum products, but the low concentrations imply contributions to the TPAH at the 0 m stations of Sleepy Bay and Northwest Bay that are less than 100 ng/g, and are questionable at that.

The low TPAH concentrations found at Olsen Bay are generally comparable with concentrations found at respective depths during the first three years after the *Exxon Valdez* oil spill. Total PAH concentrations collected earlier ranged from 6.4 ng/g at 0 m to nearly 600 ng/g at 100 m, and generally increased with depth, although a few anomalously high concentrations were found at 3 m and at 20 m (O'Clair et al., 1996).

The absence of substantial continuing contributions from intertidal *Exxon Valdez* oil to subtidal sediments is consistent with the long term persistence of the oil in the intertidal. Detection of *Exxon Valdez* oil in shallow subtidal sediments a decade after the spill would imply continuing transport of the oil from the intertidal, which would eventually deplete the intertidal oil reservoir. The fact that oil is not readily detectable in shallow subtidal sediments, especially those adjacent to beaches that remain visibly oiled (as at Northwest Bay), corroborates the persistence of the oil in the intertidal. The mechanism of transport of oil to the subtidal probably involved attachment to particles of fine sand or clay, a product of the intensive cleaning or natural processes in the 1989-01 time period. Intertidal oil present in 2001 may still have releases into the water, but not as readily attached to particles.

Most of the PAH found in 2001 at most of the stations sampled probably derive from the regional background (Page et al., 1996), which is most likely a combination of terrestrially eroded coals and organic-rich shales from the Kulthieth Formation east of the Copper River (Van Kooten et al., 2002; Short and Heintz, 2003), transported into PWS by the Alaska Coastal Current after hydraulic transport to the northern Gulf of Alaska (Page et al., 1996). The relative proportions of PAH of this background characteristically include the following: dominance of the naphthalenes and phenanthrenes/anthracenes, with the most abundant homologues bearing one to three (and usually two) alkyl carbon atoms; low abundances of dibenzothiophenes compared with phenanthrenes/anthracenes (usually less than 10%); prominent relative concentrations of benzo-b-fluoranthene, benzo-e-pyrene and benzo-*g,h,i*-perylene among the 5- and 6-ring PAH (as in Figure 3). Another characteristic of the regional background fingerprint is the presence of a regular series of n-alkanes from nC9 through nC34 (Short and Heintz, 2003), at total n-alkane concentrations that are about 80% the TPAH and highly correlated with the PAH (J. Short, Auke Bay Laboratory, unpublished data). This feature is not observed in the samples analysed for this project because the sample aliquots analyzed were not sufficiently great for these alkanes to be detected above MDLs, but conversely, the generally low n-alkane concentrations in comparison with the PAH is consistent with this aspect of the regional background. The presence of these n-alkanes in the regional background and their high correlation with the PAH strongly implies these hydrocarbons are sequestered in their inorganic matrix and are persistent on geologic time scales, and hence are not biologically available.

The PAH concentrations found during this study, including the highest concentrations at 100 m depth, are a very small fraction of concentrations in the intertidal where remnant oil from the *Exxon Valdez* is still present. When found intertidally in 2001 or subtidally in 1989-91, this *Exxon Valdez* remnant oil was usually present as a light oil residue, which typically contained around 3 mg oil/g sediment in 2001 (J. Short, Auke Bay Laboratory, unpublished data). Assuming a TPAH concentration in weathered oil of 2% (Sale et al., 1996), this is equivalent to a TPAH concentration of about 60 μg/g, approximately 100 times greater than the highest TPAH concentrations now found in subtidal sediments. The PAH in the remnant *Exxon Valdez* oil are also biologically available, unlike PAH associated with the regional background, the major source of PAH in the subtidal sediments. Because the four locations we sampled in western PWS during 2001 were among the most heavily contaminated by the *Exxon Valdez* oil spill, it is likely that remaining impacts from *Exxon Valdez* oil elsewhere are by now even lower. We therefore conclude that by 2001, contributions from remnant *Exxon Valdez* oil to the subtidal sediments are negligible in comparison with contributions from the regional background.

The anomalous PAH composition patterns evident at Sleepy Bay, Herring Bay, and Northwest Bay at respective depths of 0 m, 3 m and 10 m indicates contributions from other sources in addition to the regional background. The PAH patterns evident at these depths at Herring Bay and Northwest Bay (Figure 4) are inconsistent with the weathering pattern typical of *Exxon Valdez* oil (Short and Heintz, 1997), and the dominance of 5- and 6-ring PAH, together with the pattern of declining abundance with increasing alkyl substitution suggests a combustion source. The same pattern of declining PAH alkyl-homologue abundances at Sleepy Bay also preclude a substantial contribution from

Exxon Valdez oil, although the presence of naphthalenes, fluorenes and dibenzothiophenes suggests the possibility of low-level contributions from multiple sources, and the likely presence of phytane indicates one of these may be derived from petroleum. Natural and anthropogenic combustion sources have been proposed as potential inputs to PWS subtidal sediments (Page et al., 1996), and the TPAH concentrations found here of a few hundred ng/g illustrate the likely magnitude of such contributions.

Summary

We found no evidence that lingering oil in the intertidal from the *Exxon Valdez* oil spill continues to contaminate subtidal sediments in the adjacent subtidal. Concentrations of PAH in subtidal sediments are typically lower by factors of 100 or more compared with intertidal sediments that remain contaminated by *Exxon Valdez* oil. The major source of PAH in the subtidal sediments was fine-grained, organic-rich shales and coals eroded from terrestrial deposits east of PWS, although inputs from other unknown sources, likely related to combustion processes, were occasionally observed. Unlike PAH in *Exxon Valdez* oil, PAH in the subtidal sediments are probably not biologically available, because they are sequestered within refractory inorganic (i.e. shale) or organic (coal) matrixes.

Literature Cited

Brodersen, C., J. Short, L. Holland, M. Carls, J. Pella, M. Larsen, and S. Rice. 1999. Evaluation of oil removal from beaches 8 years after the Exxon Valdez oil spill. Proceedings of the Twentysecond Arctic and Marine Oilspill Program (AMOP) Technical Seminar, Environment Canada, Ottawa, Ont. pp. 325-336.

Carls, M.G., Babcock, M.M. P.M. Harris, G.V. Irvine, J.A. Cusick, and S.D. Rice. 2001. Persistence of Oiling in Mussel Beds after the Exxon Valdez Oil Spill. Marine Environmental Research 51:167-190.

Carlson, P. R., and K. A. Kvenvolden. 1996. Tracking *Exxon Valdez* oil from beach to deepwater sediments of Prince William Sound, Alaska. Pages 109-120 in Rice, S. D., R. B. Spies, D. A. Wolfe, and B. A. Wright (eds). Proceedings of the *Exxon Valdez* Oil Spill Symposium. American Fisheries Society Symposium 18. American Fisheries Society, Bethesda, Maryland

O'Clair, C. E., J. W. .Short, and S.D. Rice. 1996. Contamination of intertidal and subtidal sediments by oil from the *Exxon Valdez* in Prince William Sound. Pages 61-93 in Rice, S. D., R. B. Spies, D. A. Wolfe, and B. A. Wright (eds). Proceedings of the *Exxon Valdez* Oil Spill Symposium. American Fisheries Society Symposium 18. American Fisheries Society, Bethesda, Maryland

Page, D. S., Boehm, P. D., Douglas, G. S., Bence, A. E., Burns, W. A. and Mankiewicz, P. J. 1996. The natural petroleum hydrocarbon background in subtidal sediments of Prince William Sound, Alaska, USA. Environ. Toxicol. Chem. 15(8):1266-128

Sale, D. M., Short, J.W., and J.C. Gibeaut. 1996. Nearshore transport of hydrocarbons and sediments after the *Exxon Valdez* oil spill. Final Report to the Exxon Valdez Trustee Council.

Short, J. W. and Heintz, R. A. 2003. Normal alkanes and the unresolved complex mixture as diagnostic indicators of hydrocarbon source contributions to marine sediments of the northern Gulf of Alaska, in *Proceedings of the Twentysixth Arctic and Marine Oilspill Program (AMOP) Technical Seminar*, Environment Canada, Ottawa, ON, Canada.

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 PAH weathering. Environmental Science and Technology 31: 2375-2384.

Short, J.W., D.M. Sale, and J.C. Gibeaut. 1996a. Nearshore transport of hydrocarbons and sediments after the *Exxon Valdez* oil spill. Pages 40-60 in Rice, S. D., R. B. Spies, D. A. Wolfe, and B. A. Wright (eds). Proceedings of the *Exxon Valdez* Oil Spill Symposium. American Fisheries Society Symposium 18. American Fisheries Society, Bethesda, Maryland.

Short, J. W.; Jackson, T. J.; Larsen, M. L.; Wade, T. L. 1996b. In *Exxon Valdez Oil Spill Symposium*; Rice, S. D., Spies, R. B., Wolfe, D. A., Wright, B. A., Eds.; American Fisheries Society Symposium 18: Bethesda, Maryland,1996 ; pp 140–148.

Short, J. W., M. R. Lindeberg, P. M. Harris, J. M. Maselko, J. J. Pella and S. D. Rice. 2004. An estimate of oil persisting on beaches of Prince William Sound, 12 years after the *Exxon Valdez* oil spill. Environmental Science and Technology (in press).

Short, J.W., K.A. Kvenvolden, P.R. Carlson, F.D. Hostettler, R.J. Rosenbauer, and B.A. Wright, 1999. Natural Hydrocarbon Background in Benthic Sediments of Prince William Sound, Alaska: Oil vs Coal. Environmental Toxicology and Chemistry 33(1):34-42.

Van Kooten, G. W., Short, J. W., and Kolak, J. J. 2002. Low maturity Kulthieth Formation coal: A possible source of polycyclic aromatic hydrocarbons in benthic sediment of the northern Gulf of Alaska. Environmental Forensics 3:227-242.

Wolfe, D. A.; Galt, J. A.; Watabayashi, G.; Short, J.; O'Claire, C.; Rice, S.; Michel, J.; Payne, J. R.; Braddock, J.; Hanna, S.; Sale, D. 1994. The Fate of the Oil Spilled from the *Exxon Valdez*. Environmental Science and Technology 28:561A-568A.

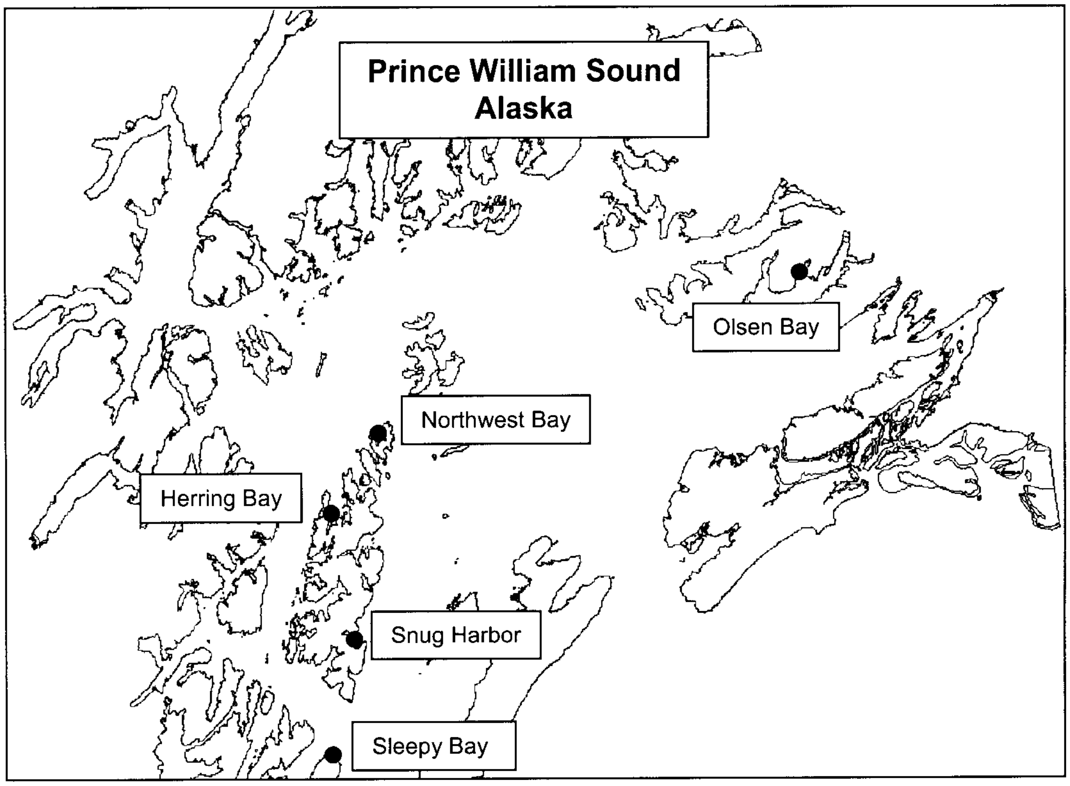


Figure 1. Sampling locations in Prince William Sound, Alaska, where sediment samples were collected along transects from 0 m to 100 m subtidal depth.

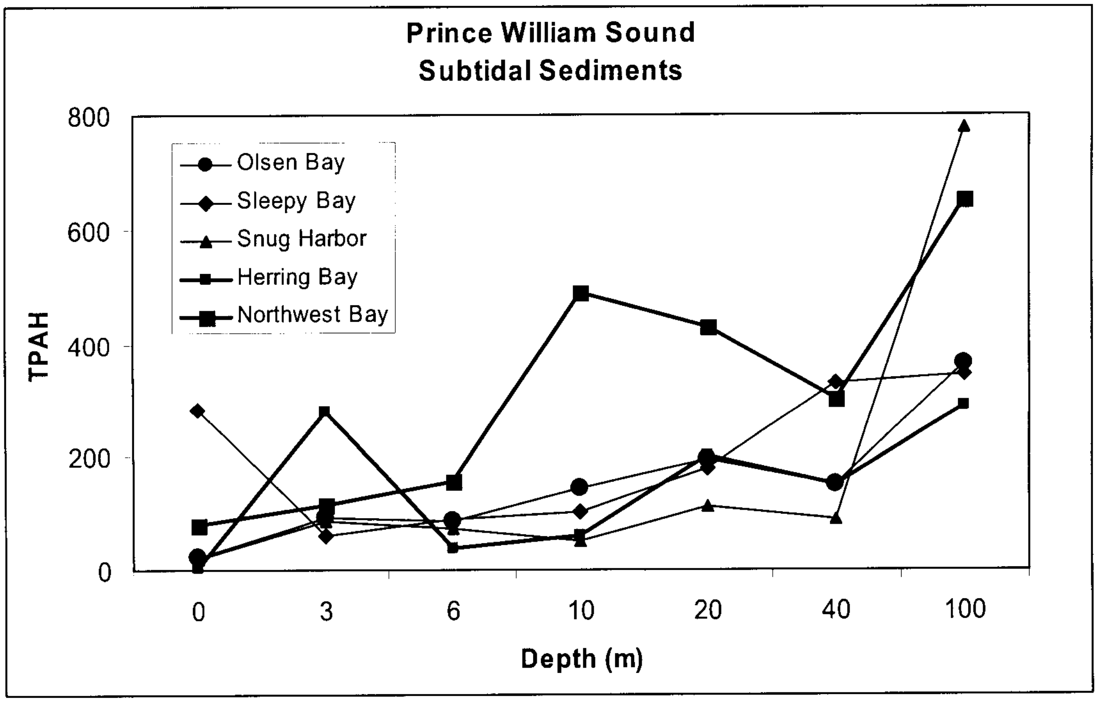


Figure 2. Concentrations of total polycyclic aromatic hydrocarbons (TPAH) in benthic sediment samples from 0 m to 100 m depths at each of the five sampling locations of this study.

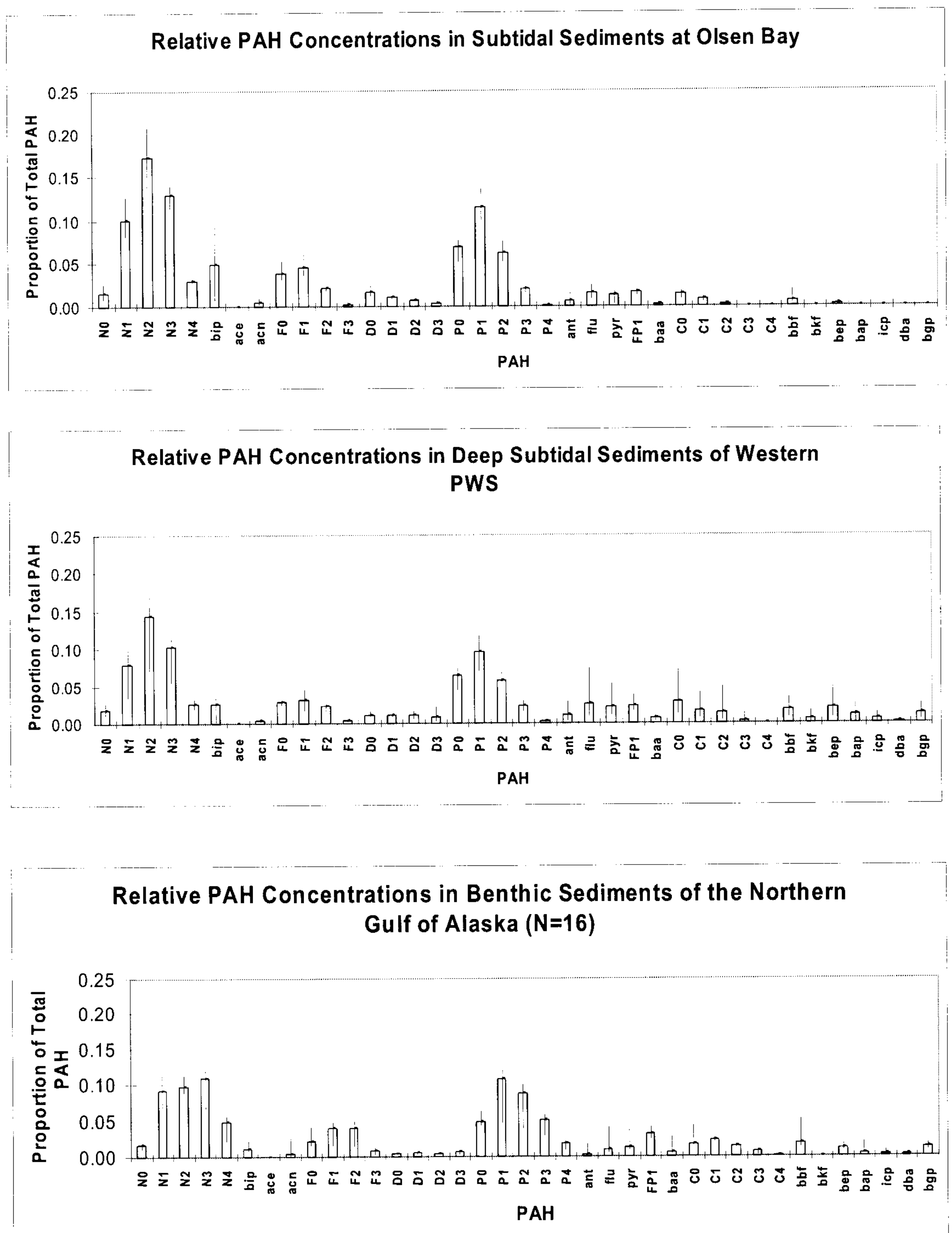
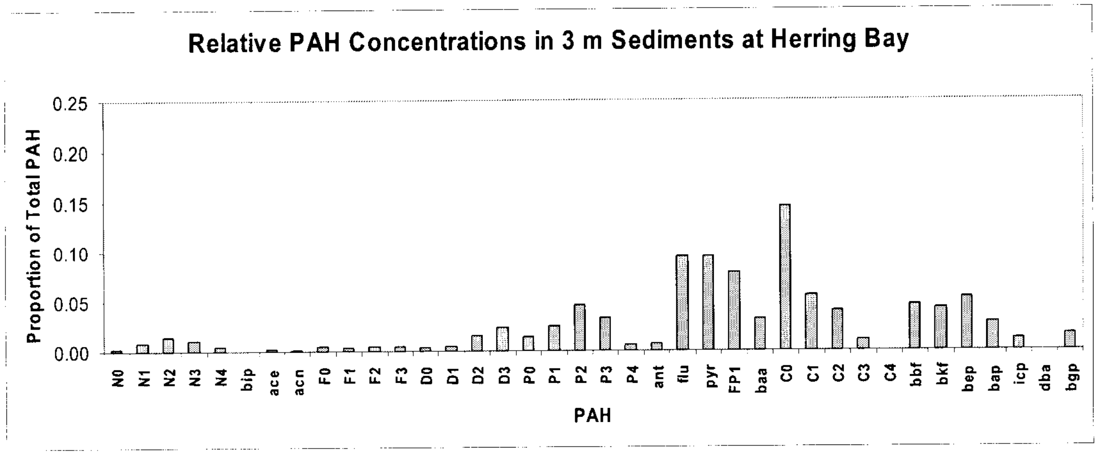


Figure 3. Relative concentrations of polycyclic aromatic hydrocarbons (PAH) in benthic sediment samples from (A) Olsen Bay, all depths, (B) station depths from 20 m to 100 m at Herring Bay, Northwest Bay, Sleepy Bay and Snug Harbor, and (C) sediments from 106 m - 308 m depths in the northern Gulf of Alaska (from Short et al., 1999). Thin vertical bars indicate ranges of relative PAH proportions.



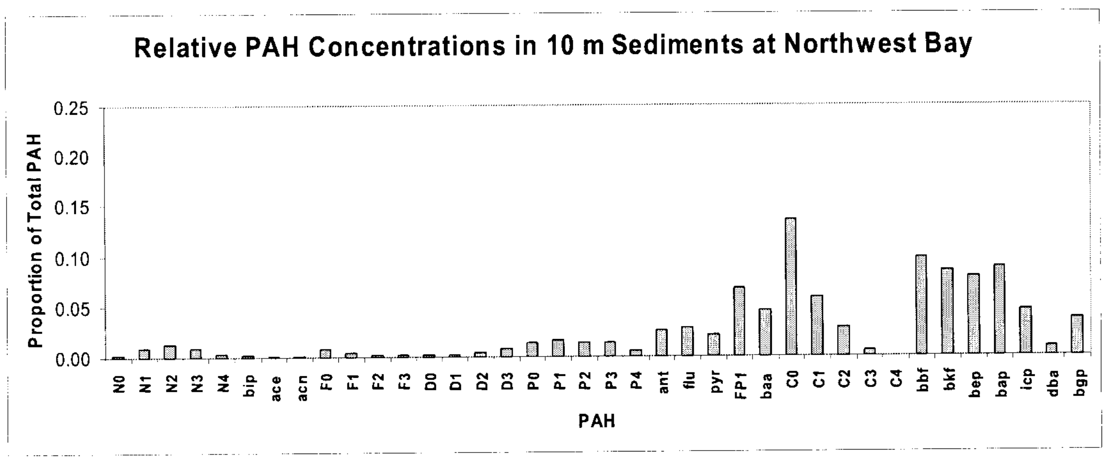


Figure 4. Relative concentrations of polycyclic aromatic hydrocarbons (PAH) in benthic sediment samples from (A) Herring Bay, 3 m and (B) Northwest Bay, 10 m.

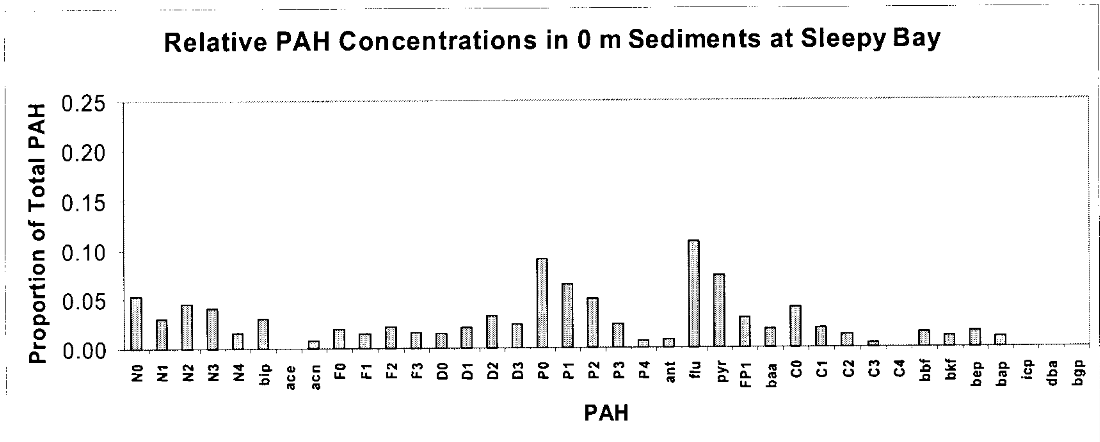


Figure 5. Relative concentrations of polycyclic aromatic hydrocarbons (PAH) in benthic sediment samples from Sleepy Bay, 0 m.