PART II

Identification of *Exxon Valdez* Oil in Sediments and Tissues from Prince William Sound and the Northwestern Gulf of Alaska Based on PAH Weathering

Identification of *Exxon Valdez* Oil in Sediments and Tissues from Prince William Sound and the Northwestern Gulf of Alaska Based on PAH Weathering

JEFFREY W. SHORT* AND RON A. HEINTZ

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

Keywords: petroleum; weathering; *Exxon Valdez*; PAH; fingerprinting; oil spill

November 1996

We used a first-order loss-rate kinetic model of polynuclear aromatic hydrocarbon (PAH) weathering to evaluate 7767 environmental samples collected for the *Exxon Valdez* oil spill (EVOS) of March 1989 for the presence of spilled oil. The model was developed from two successive experiments with gravel coated with Alaska North Slope crude oil and washed for 6 months. The 14 most persistent PAH of 44 analyzed by GC/MS were included in the PAH-weathering model. Parameters include loss-rate constants related to the energy required for PAH to escape from petroleum through the Arrhenius equation, and a quantitative index of weathering. The model accounts for 91% of the temporal variability of modeled PAH concentrations; the remaining variability is ascribed to relatively small interferences of tetramethylnaphthalenes and di- and trimethylfluorenes.

We applied the weathering model to analytical results from field samples collected for the EVOS by comparing the fit of model-predicted versus measured PAH concentrations, with a probability distribution of fits derived from the experimental weathering results. Only 1541 field samples contained sufficient PAH for valid application of the model; three-fourths fit the model at $\alpha \geq 0.01$ type I error, 9% fit an alternate model characterized by the absence of weathering, 17% fit neither model, and a few fit both models. The 1164 total samples that fit the weathering model account for 86% of all the PAH concentrations detected in all 7767 samples. We conclude that first-order loss-rate kinetics accounts for PAH weathering in the laboratory and for the dominant PAH weathering processes in the EVOS, and that the rate of weathering is determined mainly by the ratio of surface area to volume of petroleum in the environment.

Introduction

Identification of a petroleum pollution source based on chemical analysis of environmental samples is complicated by time-varying compositional changes (weathering) following introduction into the environment. The effects of weathering processes such as evaporation, dissolution, microbial degradation, and photo-oxidation that cause the composition of petroleum to vary in the environment seem difficult to predict because they are sensitive to a plethora of varying environmental conditions. Compositional changes caused by weathering have, therefore, constrained identification methods to those that are based on stable and persistent parameters derived from chemical analysis of environmental samples (*1*), or that explicitly account for the compositional changes, e.g. by comparison of analytical results from environmental samples with results from petroleum samples that have been artificially weathered to varying degrees (*2*).

The above weathering processes imply first-order (FO) loss-rate (LR) kinetics for the disappearance of polynuclear aromatic hydrocarbons (PAH) from petroleum. Evaporation and dissolution involve FOLR kinetics explicitly, with rate constants determined by the enthalpy of vaporization through the Arrhenius equation. Similar FOLR disappearance kinetics have recently been demonstrated for microbial oxidation of petroleum PAH (*3*), where PAH dissolution is promoted by biosurfactants secreted by microbes (*4*) that increase the effective surface area of the petroleum phase. The disappearance kinetics that characterize dissolution and microbial oxidation are, therefore, probably similar to evaporation qualitatively, and the distribution pattern of PAH that results from the combined effects of these processes characteristically include preferential losses of PAH that have lower molecular weight and

contain fewer alkyl substituents (*1*, *5–9*). In contrast, the kinetics of PAH photo-oxidation may be second order and autocatalytic (*10*). Photo-oxidation of PAH probably begins with photolytic generation of mainly benzylic free radicals (*11*), which would result in preferential losses of PAH that contain more numerous alkyl substituents. However, at low ambient light fluxes PAH photo-oxidation rates may be small compared with rates for evaporation and dissolution.

If the dominant PAH weathering processes include evaporation, dissolution, and microbial oxidation, but not photo-oxidation, then similar overall FOLR kinetics for PAH disappearance from petroleum result. Although the absolute concentrations of PAH constituents of petroleum may depend in a complicated way on the environmental history of a sample, the similar FOLR kinetics imposes constraints on relative PAH concentrations in weathered samples of petroleum. These constraints may provide a basis for identification of a major petroleum source in environmental samples if alternative sources can be distinguished with confidence.

We report here our experimental validation of a FOLR kinetic model to describe the disappearance of PAH from gravel coated with experimentally-weathered petroleum, and our subsequent evaluation of this model as a basis for identifying petroleum spilled from the T/V *Exxon Valdez* in environmental samples of sediments and tissues. We measured PAH concentration changes of the petroleum over 6 months of exposure to flowing, intermittently brackish water, and fit the results to a FOLR kinetic model. We developed a least-squares parameter estimation procedure for the model; these parameters include relative kinetic constants for disappearance of each of 14 PAH analytes monitored, and a metric for each

gravel sample collected, indicating the extent of weathering. We used these results to develop a probability-based method to distinguish between PAH derived from petroleum released into the area affected by the T/V *Exxon Valdez* oil spill (EVOS) of 24 March 1989 and PAH in Prince William Sound (PWS) and the northern Gulf of Alaska from alternative sources.

The EVOS provided a unique opportunity to evaluate the petroleum identification procedure we derived. Photo-oxidation was probably suppressed by the low sun angle and persistent cloud cover characteristic of the affected area, compared with more temperate regions. The EVOS area was nearly pristine before the spill (*12, 13*), and PAHs from natural sources were mainly confined to subtidal sediments in that area (*14–16*), so PAH interferences from sources other than the EVOS were often minimal in environmental samples. Moreover, the PAH pattern that characterizes the dominant natural source of PAH in marine sediments of PWS is temporally invariant and thus apparently not subject to weathering (*13*). This natural and stable PAH source provides an alternative pattern that may be used to evaluate the discriminating power of our identification method. Finally, the number of environmental samples collected and analyzed by consistent GC/MS methods (*17*) was exceptionally large (7767 analyses), and the analytical methods and spilled petroleum are directly comparable with those used for the weathering experiment we performed to evaluate the kinetic model. Therefore, the EVOS may be considered an especially large-scale field test for our kinetic model.

Model Description and Parameter Estimation

Suppose the rate of loss to the environment of a PAH (denoted as P) dissolved in petroleum follows FOLR kinetics, so that

$$-\frac{\mathrm{d}[P]}{\mathrm{d}t} = k(t)[P] \tag{1}$$

The time dependence of the LR constant, $k(t)$, derives from the variable exposure conditions of the petroleum in the environment. Writing $k(t)$ as $k\,f(t)$ and integrating eq 1 gives

$$\ln\left(\frac{[P]_0}{[P]}\right) = k\int_0^t f(t)\mathrm{d}t = kw \tag{2}$$

where the value of the integral in eq 2 is indicated by a weathering parameter, w, which summarizes the exposure history of the petroleum volume element sampled.

Equation 2 may be simultaneously applied to J different PAH in each of I different samples as

$$\ln\left(\frac{[P_{ij}]_0}{[P_{ij}]}\right) = k_j w_i \tag{3}$$

where j specifies the PAH and i specifies the sample. Note that each k_j is the same for all samples, and that each w_i is the same for all the PAH in a sample. If the initial concentrations $[P_{ij}]_0$ in an environmental sample are known, the parameters k_j and w_i may be estimated from measurements of the $[P_{ij}]$ in the samples (see below).

The initial PAH concentrations $[P_{ij}]_0$ in an environmental sample may be estimated from the initial amount and composition of the petroleum in the sample. The initial amount

of petroleum introduced into unit mass of the sampled environmental matrix can be determined if there are temporally invariant analytes contained in the petroleum, denoted here as j^+ (and assuming the analyte set is restricted to PAH). Denoting the proportion of the jth PAH in the unweathered petroleum as π_j, then the concentration c_i of unweathered petroleum originally introduced into the ith sample is determined as the ratio of $[P_{ij+}]$ (= $[P_{ij+}]_0$) to π_{j+}. The mean of these ratios may be used as an estimator of c_i,

$$\hat{c}_i = \frac{1}{n_{j+}} \sum_{j+ \in J} \frac{[P_{ij+}]}{\pi_{j+}} \tag{4}$$

where n_{j+} indicates the number of temporally invariant PAH included in the sum. Because $[P_{ij}]_0 = \hat{c}_i \pi_j$, eq 3 may be expressed as

$$d_{ij} = \ln \frac{\hat{c}_i \pi_j}{[P_{ij}]} = k_j w_i \tag{3a}$$

Note that the quantity indicated by d_{ij} may be determined entirely from measurements of PAH concentrations in the unweathered petroleum and in an environmental sample.

Error minimization of $e_{ij} = d_{ij} - \hat{k}_j \hat{w}_i$ by least-squares leads to

$$\hat{k}_j = \frac{\sum_{i=1}^{I} \hat{w}_i d_{ij}}{\sum_{i=1}^{I} \hat{w}_i^2}; \qquad \hat{w}_i = \frac{\sum_{j=1}^{J} \hat{k}_j d_{ij}}{\sum_{j=1}^{J} \hat{k}_j^2} \tag{5}$$

It can be shown that eqs 5 imply that the vector $\hat{\mathbf{k}}$ of elements $\hat{k}_j$ may be identified with the dominant eigenvector of the matrix **D'D**, where the elements of **D** are d_{ij} (see appendix). Imposition of the normalization condition $\sum \hat{k}_j^2 = 1$ removes the indeterminacy implied by eq 3, so that $\hat{\mathbf{k}}$ may be identified with the first principal component of **D'D**, with corresponding eigenvalue $\lambda_1 = \sum \hat{w}_i^2$. Thus, a principal component analysis of **D'D** provides least-square estimates of k_j, and the $\hat{w}_i$ may then be calculated from eqs 5 and the d_{ij}. In this context, the ratio of the dominant eigenvalue λ_1 to the sum of the eigenvalues (i.e. the trace of **D'D**) may be interpreted as the proportion of data variability explained by the model. Also, $\exp\pm(\sum_{j>1}^{J} \lambda_j)/IJ)^{1/2}$ provides a measure of the root-mean-square fit of predicted and observed PAH concentrations.

The distribution of $\sum_j e_j^2$ provides a basis for evaluating the probability that PAH in an environmental sample is consistent with weathered *Exxon Valdez* oil (EVO) as the PAH source. Given $\hat{\mathbf{k}}$ and $\hat{w}_i$ derived from PAH measurements of experimentally weathered EVO samples, predicted PAH concentrations may be calculated for sample i by eq 3a. The agreement of predicted and measured PAH concentrations in the ith sample may be expressed as the mean of the squared differences of logarithms of measured PAH and predicted PAH, or mean square error (MSE_i):

$$\frac{1}{J}\sum_{j=1}^{J}\left[\ln\left(\frac{[\hat{P}]_{ij}}{[P]_{ij}}\right)\right]^2 \quad = \quad \frac{1}{J}\sum_{j=1}^{J}\left(d_{ij} - \hat{d}_{ij}\right)^2 \quad = \quad MSE_i \qquad (6)$$

The quantity $MSE_{i'}$ may be similarly calculated for PAH of uncertain origin in an environmental sample i', and compared with the MSE_i distribution developed by the bootstrap method applied to PAH measurements of experimentally weathered EVO samples. An

estimate of the probability that the PAH in the environmental sample i' is consistent with the experimentally weathered petroleum may then be made on the basis of this comparison.

If PAH concentrations can be predicted on the basis of alternative PAH sources, the same approach may be used to estimate the probability that PAH in an environmental sample is consistent with the alternative PAH source, provided the alternative mean square error distribution is available. In particular, if the alternative PAH source is not subject to weathering, so that relative PAH concentrations do not change with time, then PAH proportions of the sum of the PAH measured provides a basis for prediction. Equation 6 may be used to calculate the alternative mean square error MSE_i', which can be compared with the MSE_i' distribution calculated by the bootstrap method from known samples containing PAH from only the alternative source. An estimate of the probability that the PAH in the environmental sample is consistent with the alternative PAH source may then be made on the basis of this comparison.

Methods

Petroleum Sources and Composition. The petroleum spilled from the T/V *Exxon Valdez* was produced from the Alaskan North Slope (ANS) oil fields in 1989, and the petroleum used in the weathering experiment was produced from the same fields in 1992 and 1993. The PAH analytes we considered are listed in Table 1, with proportions of these PAH analytes in unweathered petroleum from the cargo of the T/V *Exxon Valdez* as determined by the hydrocarbon analysis methods used herein (*17*). Also in Table 1 are the PAH analytes we selected for modeling, based on analyte persistence above method detection limits (MDL; see

below) during the experimental weathering period. The proportions of selected PAH in unweathered EVO given in Table 1 are taken as the π_j in developing the weathering model.

Petroleum Weathering Experiments. Petroleum produced from the ANS was experimentally weathered by continuously washing petroleum-coated gravel for 6 months. The petroleum was heated at 70 °C overnight to 80% initial mass to remove volatiles, then sprayed onto tumbling gravel. Four different loadings of petroleum on gravel were prepared in 1992 and three in 1993, when the weathering experiment was repeated. The 1992 loadings were 55.2, 622, 3130, and 4510 μg petroleum/g gravel, and were 281, 717, and 2450 μg petroleum/g gravel in 1993. Each 11-kg preparation of petroleum-coated gravel was weathered by continuously washing it in a polyvinyl chloride (PVC) tube with alternating fresh and 30‰ seawater, switching every 6 h, at flow rates of 150 mL/min. Water temperatures ranged from 12 °C at the beginning of the weathering period to 2 °C in midwinter. Further details on the methods used to determine petroleum loadings on gravels, gravel preparation, and weathering apparatus and procedures have been presented previously (*18*).

Gravel samples were composited for PAH analysis from 8–15 replicate preparations of each petroleum loading. At each sampling, equal numbers of gravel pieces were removed from each replicate PVC tube, mixed, and stored at -20 °C until analysis. Gravel samples were collected 5, 62, 90, and 180 days after the columns were filled in 1992, and 3, 41, 68, and 181 days after the columns were filled in 1993. Duplicate composite samples were collected at day 62 in 1992; otherwise, single composite samples were collected. Of the 32

composite samples, 6 were not used to develop the weathering model because one or more of the PAH analytes selected for modeling were below MDL.

The PAH content of samples was determined by a GC/MS method. The analytes include unsubstituted and alkyl-substituted homologues of 2- to 4-ring PAH, and dibenzothiophene homologues (Table 1). PAH were extracted with dichloromethane, purified by alumina/silica gel column chromatography followed by size-exclusion high-performance liquid chromatography. Purified PAH were separated by GC and measured by MS operated in the selected ion monitoring mode. Concentrations of PAH in the dichloromethane extracts were determined by the internal standard method based on a suite of deuterated-PAH internal standards. Four quality control samples were analyzed with each batch of 12 samples, including 2 reference samples, a method blank, and a method blank spiked with certified hydrocarbon standards obtained from the National Institute of Standards and Technology (NIST). Method detection limits of hydrocarbon analytes were determined experimentally (*19*), and were generally 1 ng/g. At the Auke Bay Laboratory (ABL), the accuracy of this analysis is better than about ±15% based on comparison with NIST values, and precision expressed as coefficient of variation is about 25% for the analytes in the weathering model, based on reference sample results (see data analysis). Additional details of the method used have been presented elsewhere (*17*).

***Exxon Valdez* Oil Spill Study Area.** The EVOS introduced some 35,500 metric tons of ANS petroleum into PWS, which traveled about 750 km southwest along the Kenai peninsula and through Shelikov Strait before dispersing into the northern Gulf of Alaska, oiling about 1750 km of shoreline along the way (*20*). The path followed by the spilled

petroleum conformed with the Alaska Coastal Current (ACC) (Figure 1), which flushes PWS through Hinchinbrook Entrance and exits through Montague Strait (*21*). Sea-surface temperatures of the affected area typically range from near 0 to 16 °C annually. The ACC transports sediments burdened with PAH from natural sources into PWS (*16*). The pattern of relative PAH concentrations characteristic of sediments transported by the ACC into PWS has been consistently found at intertidal stations near Hinchinbrook Entrance (*13*, *16*), and has often been found subtidally within PWS (*13–16*). At Constantine Harbor off Hinchinbrook Entrance, concentrations of selected PAH in intertidal sediments have been constant at least since 1977, and probably much longer (*13*).

After the EVOS, 3433 samples of sediments, 2150 samples of mussels (*Mytilus trossulus*), and 2184 samples of other tissues were collected and analyzed for PAH to support the natural resource damage assessment efforts of the state and federal governments. ABL staff collected most of these sediments and mussels, using dichloromethane-rinsed apparatus, and stored them in pre-cleaned glass jars fitted with polytetrafluoroethane cap-liners at −20 °C until analysis. Procedures used to collect the remaining samples were usually similar. The PAH analysis methods used (*17*) were identical with those summarized above for the petroleum weathering experiment.

Database Archive. All of the hydrocarbon analysis results for this report are contained in the EVOS of 1989 State/Federal Trustee Council Hydrocarbon Database (EVTHD) at the ABL and available on internet at www.xxx.xxx.xxx.gov. Results of all sampling information and hydrocarbon analyses were entered into a data repository before being reviewed by the principal investigators responsible for the sample collections. Data in

this repository, named PWSOIL, were transferred to EVTHD after principal investigators reviewed database sampling information and analytical results for consistence with their project records. Only data for environmental samples were transferred; experimentally manipulated samples, method blanks, spiked samples, samples with incomplete information, and duplicate analyses were not included with EVTHD. We treat hydrocarbon results incorporated into EVTHD below MDL as zero for the purposes of this report.

Data Analysis

Petroleum Weathering Experiments. We could not apply the weathering model to all the PAH initially present in EVO because progressively more PAH were below MDLs as the petroleum weathered during the weathering experiment. We therefore applied the weathering model to the most broadly persistent PAH selected from the five most prominent PAH-homologue groups in EVO: naphthalenes, fluorenes, dibenzothiophenes, phenanthrenes, and chrysenes. Nine of the selected PAH were present above MDL in all 32 samples collected during the petroleum weathering experiments and provided 288 observations of PAH for the model. Another 5 PAH were simultaneously present in 26 of the 32 samples, and the 14 PAH simultaneously present in these 26 samples provided 364 observations of PAH. This combination of 14 PAH in 26 samples provided the maximum number of simultaneous PAH observations possible. We therefore applied the weathering model to the $J = 14$ PAH (identified in Table 1) simultaneously present in $I = 26$ of the 32 samples collected during the petroleum weathering experiment.

We calculated the initial oil concentration parameter ($\hat{c}_i$) for each sample based on the modeled chrysene homologues, which are persistent in weathered crude oil (*5*). The sum of chrysene, C-1 chrysene, and C-2 chrysene concentrations did not change significantly with time in 1992 or 1993 (repeated measures ANOVA; $P > 0.05$), therefore these homologues were used to estimate $\hat{c}_i$ from eq 4. Gravimetric determinations of petroleum initially applied to the gravel used in the weathering experiments were linearly related to $\hat{c}_i$ ($r^2 = 0.86$, $P = 0.02$), and were about 60% lower after correcting for volatility losses. We calculated $\hat{w}_i$ for the 26 samples and relative $\hat{k}_j$ for the 14 PAH from the 364 PAH observations obtained from the petroleum weathering experiment by principal component analysis of the matrix $\mathbf{D'D}$, after transformation of PAH observations to the matrix elements d_{ij} of $\mathbf{D}$.

Bootstrapped distributions of the MSE_i and $\hat{k}_j$ were simultaneously constructed by Monte Carlo simulation. One of the 26 samples was randomly removed, and a new matrix $\mathbf{D^*}$ was calculated from the PAH observations of 26 random selections with replacement from the 25 remaining samples. New LR constants were calculated by finding the eigenvector $\hat{\mathbf{k}}^*$ of $\mathbf{D^{*\prime}D^*}$, and the MSE_i was calculated for the removed sample from eq 6 using $\hat{\mathbf{k}}^*$. This process was repeated 500 times, and the 14 $\hat{k}_j^*$ and MSE_i were recorded for each iteration. The distribution of the 14 $\hat{k}_j^*$ is presented as the range of the central 95% of the bootstrap results for each $\hat{k}_j$. The frequency distribution of MSE_i is used to estimate the probability that PAH in environmental samples are consistent with PAH from weathered ANS petroleum.

Characterization of PAH in Sediments from Natural Sources. We based our model of relative PAH concentrations that characterize the natural PAH source on samples collected from Constantine Harbor. The relative PAH concentration pattern and its error

distribution are derived from 15 intertidal sediment samples collected during six samplings in 1989 and 1990. Because total PAH (TPAH, i.e., the sum of the PAH analyzed) did not vary significantly (ANOVA, $P > 0.23$) among these samples (*13*), the sample with the median value of TPAH was arbitrarily selected as representative of the characteristic PAH pattern, and an error distribution for this pattern was generated by comparing the remaining samples with this representative sample as follows. The same 14 PAHs used in the weathering model were each converted to proportions by dividing each PAH concentration for a sample by the TPAH of the sample. We denote these PAH proportions in the median sample $\tilde{i}$ and a different sample i as $p_{\tilde{i}j}$, and p_{ij}, respectively, and calculate MSE_i' for discrepancies among these proportions as

$$\frac{1}{14}\sum_{j=1}^{J}\left(\ln[p_{\tilde{i}j} \,/\, p_{ij}]\right)^2 \quad = \quad MSE_i' \tag{7}$$

A bootstrapped distribution for MSE_i' was constructed by iterating the following procedure 500 times. One of the 15 samples from Constantine Harbor was randomly removed, and the remaining 14 samples were sampled with replacement 15 times. The sample of the resulting set with the median TPAH value was selected as a new representative of the characteristic PAH pattern. The PAH proportions of TPAH for the new median sample and for the sample initially removed were respectively denoted as $p^*_{\tilde{i}j}$ and p^*_{ij}, and these proportions were used in eq 7 to calculate a new observation of MSE_i'. The collection of 500 such observations was taken as the empirical error distribution for MSE_i'.

Hypothesis Testing. We used the bootstrapped error distributions for the weathering model and the natural sediment PAH source as the basis for distinguishing PAH sources in environmental samples. All the environmental samples with concentrations of the 14 selected PAH above MDL were consecutively fit to both models, and *MSE* and *MSE'* were calculated for each sample by eq 6 and 7 respectively. The probability that the source of PAH in the *k*th sample was consistent with ANS petroleum, denoted here as $Pr_{oil}(k)$, was determined by subtracting from 1 the percentile of values $\leq MSE_k$ in the cumulative frequency distribution of MSE_i for the weathering model. The null hypothesis that the PAH pattern in the *k*th sample was consistent with ANS petroleum was rejected when $Pr_{oil}(k) < \alpha$, where α specifies the probability of type I error. Similarly, $Pr_{ns}(k)$, the probability that the pattern of PAH in the *k*th sample was consistent with the natural sediment source, was derived by comparing MSE_k' to the cumulative frequency distribution of MSE_i'. The null hypothesis that the *k*th sample is consistent with the natural sediment pattern was rejected when $Pr_{ns}(k) < \alpha$. Note that the two models must be considered as separate alternatives rather than simultaneously [(as by, *e.g.*, SIMCA (*22*) methods], because the models are not isolinear and, therefore, cannot be combined into the same principal component matrix.

Results

Weathering Model Parameters. The eigenvector of $\hat{k}_j$ calculated as the first principal component of the matrix **D'D** accounts for 86% of the total variability in the PAH data from the petroleum weathering experiments. The mean unexplained variability per sample and per analyte is 0.161, indicating that most of the PAH values predicted by the weathering model

are within −33% and +49% of observed values from the petroleum weathering experiments. Three analytes, C4-naphthalene and C2- and C3-fluorene, together account for 73% of the magnitude of the second principal component of the **D'D** matrix, and the second principal component accounts for an additional 8% of the total variability in the PAH data from the petroleum weathering experiments. The mean variability of logarithmically-transformed data per sample and per analyte that is not explained by the first two principal components is 0.069, which implies a mean coefficient of variation of 26% for the untransformed data, consistent with analytical precision. The first two principal components, therefore, account for all the data variability except analytical error; the first component accounts for 91% [i.e. (86/0.94)%] of the explainable data variation. Thus, the second principal component summarizes the discrepancies between the weathering model and the data after discounting variability due to analytical error.

The values of $\hat{k}_j$ increase with decreasing alkyl-substitution and number of aromatic rings (Table 2). The largest values of $\hat{k}_j$ are for C3-naphthalene, C1-dibenzothiophene, and C1-phenanthrene, whereas the smallest values are for C3- and C4-phenanthrene and the three chrysene homologues. The proximity to zero of the $\hat{k}_j$ of these latter five analytes indicates that the duration of the weathering experiments was insufficient for appreciable weathering loss to occur. Also in Table 2 are the ranges for the most central 95% of the bootstrap estimates of the $\hat{k}_j$ as a measure of dispersion of the estimates. This dispersion is proportionally least for those analytes that changed most in concentration during the weathering experiments.

The values for $\hat{w}_i$ increase linearly with time during the petroleum weathering experiments, and increase faster at lower petroleum loadings (Figure 2). In the 1992 experiment, these linear trends are significant ($P < 0.027$) for all but the lowest petroleum loading ($P < 0.09$). Values for $\hat{w}_i$ range above 7 for the most heavily weathered, lowest petroleum-loaded gravel at 85 days exposure but do not exceed 3.5 for the most heavily loaded gravel at 175 days. The weathering rate, $d\hat{w}_i/dt$, increases linearly with decreasing petroleum loading ($r^2 = 0.98$; $P < 0.01$) in the 1992 weathering experiments. Similar trends occur in the 1993 experiment, but the more limited 1993 data preclude a meaningful statistical summary.

The distribution of MSE_i derived from the fit of the bootstrapped iterations of the weathering model and the PAH data from the petroleum weathering experiments (see eq 6) is strongly leptokurtic (Figure 3). The MSE_i ranges from 0.0086 to 1.47, with a median of 0.145. The 95th and 99th percentiles occur at $MSE_i = 0.57$ and 0.98, respectively, and the latter value is used below to evaluate samples from the EVOS. This corresponds with accepting a type I error probability of 0.01 when evaluating the null hypothesis that PAH patterns of environmental samples are consistent with weathered EVO. A comparison of observed and predicted PAH proportions that correspond with the median MSE_i for a weathered ($\hat{w}_i = 3.95$) example is depicted in Figure 4B, where the PAH proportions of the unweathered EVO are also presented for further comparison (Figure 4A).

Characterization of PAH in Sediments from Constantine Harbor. The PAH of Constantine Harbor intertidal sediments are proportionally lower in naphthalenes and dibenzothiophenes, and higher in phenanthrenes and chrysenes compared with EVO (Figure

4C). The distribution of MSE_i' derived from application of eq 7 to the bootstrapped iterations of the Constantine Harbor samples is less strongly leptokurtic than the MSE_i distribution of weathered EVO (compare Figures 3 and 5). The MSE_i' of the logarithmically-transformed PAH data ranges from 0.0066 to 0.37, with a median of 0.056, and this median is equivalent to a coefficient of variation of 24% for untransformed data, consistent with analytical precision. The 95th and 99th percentiles occur at $MSE_i' = 0.28$ and 0.34, respectively, and the latter value is used below to evaluate samples from the EVOS. This corresponds with accepting a type I error probability of 0.01 when evaluating the null hypothesis that PAH patterns in environmental samples are consistent with PAH from the natural source.

Classification of Sediment Samples from the EVOS. The results of our classification procedure indicate that although EVO did not contaminate most of the EVOS sediment samples collected, EVO was the source of most of the PAH detected. Concentrations of the 14 PAH included in the weathering model are above MDL in 996 of the 3433 sediment samples analyzed for the EVOS. Of these 996 samples, 618 have $MSE_k < 0.98$ and $MSE_k' > 0.34$, which we accepted as consistent with weathered EVO (Figure 6A). The sum of all the PAH concentrations detected above MDL in these 618 samples is more than 86% of the total sum of all PAH concentrations detected in all the sediment samples analyzed.

The sediment samples we classified as contaminated by EVO may also contain PAH from other contamination sources. The median MSE_k of the 618 EVO-contaminated samples was 0.34, or more than twice the median MSE_i derived from the bootstrap distribution of the petroleum weathering experiments. The larger value of the median MSE_k may be caused by PAH from other sources that alter the relative PAH proportions of sediment samples and

consequently fit the weathering model less well. This is most evident for samples that contain relatively low total PAH concentrations. Of the 618 EVO-contaminated samples identified, the median MSE_k for the 255 of these samples that have total PAH concentrations less than 750 ng/g (dry weight) is 0.47, compared to a median of 0.25 for the remaining 363 samples that have total PAH concentrations greater than 750 ng/g. The distribution of the MSE_ks for these 363 samples is strongly leptokurtic, and similar to the MSE_i distribution derived by bootstrapping results of the petroleum weathering experiments.

Most of the EVO-contaminated sediments we identified were collected from the inter- and shallow-subtidal within the EVOS impact area (Figure 1). Epibenthic surface depth was reported for 546 EVO-contaminated sediment samples; 93% of these were collected above 20 m subtidal depth within the EVOS area. Another 5.3% were collected from subtidal depths below 20 m within the EVOS area, and 1.7% were collected outside the EVOS area.

Of the 996 sediment samples we evaluated, 110 samples had $MSE_k > 0.98$ and $MSE_k' < 0.34$, which we accepted as consistent with PAHs derived from the natural PAH source (Figure 6A). The sum of all the PAH concentrations detected above MDL in these 110 samples is less than 0.06% of the total sum of all PAH concentrations detected in all the sediment samples analyzed. The median MSE_k' of the 110 EVO-contaminated samples was 0.17, or more than three times the median MSE_i' derived from the bootstrap distribution of the intertidal Constantine Harbor sediments. As with the EVO-contaminated sediments, this larger value of the median MSE_k' may be caused by PAH from other sources that alter the relative PAH proportions of environmental sediment samples and thereby fit the Constantine Harbor pattern less well.

Most of the sediments we identified as contaminated by PAH from the natural source were collected from deeper-subtidal depths within the EVOS impact area (Figure 1). Over 80% of these sediments collected within the EVOS area were from subtidal depths below 20 m. In contrast, most of the sediments collected east of the EVOS area that we identified as contaminated by PAH from the natural source were from the Constantine Harbor intertidal, which was used to define this PAH pattern.

Of the remaining sediment samples we evaluated, 30 fit both the weathering model and the Constantine Harbor PAH pattern, and 238 fit neither. The 30 samples that fit both patterns at the $\alpha = 0.01$ type I error rate (i.e. $MSE_k < 0.98$ and $MSE_k' < 0.34$) account for 0.02% of the total sum of all PAH concentrations detected in all the sediment samples analyzed. At $\alpha = 0.05$ type I error rate, no sample fit both patterns simultaneously. The 238 samples that fit neither PAH pattern ($MSE_k > 0.98$ and $MSE_k' > 0.34$) include 41 samples of sediment trap filtrates that were contaminated during sample collection (*23*), and 5 samples of EVO that were so diluted for analysis that C-1 dibenzothiophenes were detected just above MDL but well below concentrations predicted by the weathering model, which caused the poor fit to the weathering model. The PAH in the remaining samples of this category account for 0.31% of the total sum of all PAH concentrations detected in all the sediment samples analyzed, and may include mixtures of PAH from natural sources, the EVOS, and other, unknown sources.

Most of the 2437 other sediment samples that could not be evaluated contained low PAH concentrations. These samples could not be evaluated because 1 or more of the 14 PAH used in the weathering model were below MDL. The sum of the all the PAH concentrations

detected above MDL in these samples is 11.5% of the total sum of all PAH concentrations detected in all the sediment samples analyzed (Figure 6A). However, most of these PAH were in a few samples that were visibly contaminated with EVO but were over-diluted for PAH analysis, which resulted in concentration estimates for some of the 14 modeled PAH below sample mass-adjusted MDLs. For example, the sample mass analyzed for 19 of these samples was ≤50 mg, but the PAH concentrations detected above MDL in these samples account for 8.3% of the total sum of all PAH concentrations detected in all the sediment samples analyzed. The remaining 2418 samples account for 3.2% of the total sum of PAH concentrations detected.

Classification of Mussel Samples from the EVOS. As with sediments, the results of our classification procedure indicate that although EVO did not contaminate most of the mussel samples collected, EVO was the source of most of the PAH detected. Concentrations of the 14 PAH included in the weathering model are above MDL in 452 of the 2150 mussel samples analyzed for the EVOS. Of these mussel samples, the $MSE_k < 0.98$ and the $MSE_k' > 0.34$ for 435 samples, which we accepted as consistent with weathered EVO (Figure 6B). The sum of the all the PAH concentrations detected above MDL in these 435 samples is >84% of the total sum of all PAH concentrations detected in all the mussel samples analyzed. The median MSE_k of the 435 EVO-contaminated mussel samples was 0.25, about 75% more than the median MSE_i derived from the bootstrap distribution of the petroleum weathering experiments. As with the 363 EVO-contaminated sediment samples above, the distribution of the MSE_ks for these 435 mussel samples is strongly leptokurtic, and similar to the MSE_i distribution derived by bootstrapping results of the petroleum weathering experiments. In

contrast with sediment samples, the median MSE_k for sets of EVO-contaminated mussels varies little, regardless of the minimum PAH content of the sample set.

The smaller median MSE_k for mussels compared with sediments suggests that mussels were less subject to PAH contamination from sources other than the EVOS. In particular, the smallest MSE_k' of mussels is 0.57, which indicates that the Constantine Harbor sediment-PAH pattern is absent entirely in mussels, and also that no mussel simultaneously fit the weathering model and the Constantine Harbor pattern. However, 17 mussel samples fit neither PAH pattern ($MSE_k > 0.98$ and $MSE_k' > 0.34$), and these account for 0.95% of the total sum of all PAH concentrations detected in all the mussel samples analyzed (Figure 6B).

All but three of the EVO-contaminated mussel samples we identified were collected from within the EVOS impact area: two had ambiguous location information reported, and one was reported as collected from eastern PWS.

As with sediments, most of the 1699 other mussels that could not be evaluated contained low PAH concentrations. The sum of all the PAH concentrations detected above MDL in these samples is 15.0% of the total sum of all PAH concentrations detected in all the mussel samples analyzed (Figure 6B).

Classification of Other Tissue Samples from the EVOS. Samples of other tissues are classified as EVO-contaminated by the weathering model less frequently than sediments and mussels. Concentrations of the 14 PAH in the weathering model are above MDL in 93 of the 2184 other tissue samples analyzed for the EVOS. Of these 93 samples, the $MSE_k <$ 0.98 and the $MSE_k' > 0.34$ for 80 samples, which we accepted as consistent with weathered EVO (Figure 6C). The sum of all the PAH concentrations above MDL in these 80 samples is

34% of the total sum of all PAH concentrations detected in all the other tissue samples analyzed. Most of these 80 samples were from external surfaces of oiled animals. Another 12 samples did not fit either model ($MSE_k > 0.98$ and the $MSE_k' > 0.34$) and account for another 42% of detected PAH, most (>90%) of which is due to three stomach content samples from bald eagles. One sample fit both models. The remaining 23% of detected PAH is distributed among the 2090 samples of other tissues that could not be evaluated, usually at concentrations near MDLs.

Time-Dependence of the Weathering Parameter in EVO-Contaminated Sediments and Mussels. The weathering parameter $\hat{w}_i$ was only weakly correlated with the sample collection date of sediment or mussel samples identified as EVO-contaminated by the weathering model ($r^2 = 0.045$, $P < 0.001$; Figure 7). Values of $\hat{w}_i$ ranged from near zero to >7 during each of the 6 years following the EVOS, which together with the small proportion of variation in $\hat{w}_i$ explained by the sample collection dates indicates that the effect of time on weathering rate varies considerably.

Discussion

Assessment of First-Order Weathering Kinetics for Experimentally Weathered EVO. The principal component analysis of the logarithmically-transformed PAH results from the petroleum weathering experiments indicate three factors that determine the observed PAH variability. These three factors are FOLR kinetics, analytical error, and remaining variability summarized by the second principal component, which we denote as process error. Because performance of any model is constrained by the analytical error, that error must be estimated

in order to evaluate model performance. We accept the results of laboratory analysis of reference samples as the basis for analytical error estimation because these samples include matrix effects and the large number of repetitive analyses available for these analyses records analytical variability over a period of years. Because the squared coefficient of variation for PAH in reference sample results is equivalent to the variance of logarithmically-transformed PAH results, these variation coefficients may be used to calculate the approximate analytical error variance expected. On this basis, the meaningful principal components are limited to the first two: FOLR kinetics and process error.

Comparison of the eigenvalues of the first two principal components shows that the process error component is at most a minor perturbation of the FOLR process. The process error component may indicate incorrect specification of the weathering model (i.e. PAH do not weather according to FOLR kinetics), or alternatively may indicate systematic experimental errors; its composition suggests the latter rather than the former. The three largest PAH constituents of the process error component, C4-naphthalenes and C2- and C-3 fluorenes, could be due to unknown analytical interferences, or to composition differences between EVO and ANS petroleum. The composition constants π_j in our weathering model are derived from EVO, but the ANS petroleum we used was produced 3 years after the EVOS, and may have somewhat different PAH composition due to variable contributions from different ANS oil fields (*24*). The weathering model accounts for 94% of the PAH variability when the composition constants π_j are derived from the composition of the ANS petroleum initially applied to the petroleum weathering gravels, and the remaining variability is due to

analytical error. We therefore conclude that, within the limitations imposed by analytical and systematic errors, PAH variability for the weathering experiments follows FOLR kinetics.

PAH vaporization from petroleum may be considered as an endothermic chemical reaction that involves breaking the cohesive bonds between PAH solutes and the petroleum solvent. The physical rate-limiting step (RLS) implied by FOLR kinetics is the energy required to overcome the attractive van der Waals forces between the petroleum phase and departing PAH molecules that constitute these bonds. This energy requirement is approximately equal to the enthalpy of vaporization, which is proportional to molecular surface area. The Arrhenius rate equation gives the relation between rate constant k, activation energy E_a, and temperature as $k = A \exp{-(E_a/RT)}$. The linearity of a plot of ln k versus E_a derived from observations of aqueous dissolution rates of PAH from petroleum is evidence of a similar RLS for PAH vaporization and aqueous dissolution. An Arrhenius plot ln $\hat{k}_j$ versus estimates of total molecular surface area (TSA), used here as a surrogate for enthalpies of vaporization (which are not available for PAH vaporization from petroleum), is approximately linear ($r^2 = 0.75$, $P < 0.005$; Figure 8). This linearity corroborates initial separation of PAH from the petroleum phase as the RLS, regardless of the nature of the phase receiving the PAH lost from the petroleum. This also explains why the weathering model performs equally well with EVO in subtidal sediments where the receiving phase is aqueous, and with intertidal sediments and mussels, where the receiving phase may at times be the atmosphere.

An Arrhenius plot of logarithms of FOLR constants reported for a petroleum-weathering field experiment conducted at higher temperatures [Table 3 in (*3*)] is also linear (r^2

= 0.90, $P < 0.001$; Figure 8), but has a significantly ($P < 0.001$) less slope, which implies less variability with TSA among these rate constants compared with ours. The linearity corroborates the proposed RLS, but the smaller slope is due to expected temperature effects on the rate constants implied by the Arrhenius rate equation. Differences among rate constants increase with decreasing temperature when E_a is independent of temperature, and these differences are exacerbated in this case by incipient crystallization of PAH in petroleum at temperatures near 0 °C, which would increase vaporization enthalpies of larger PAH compared with warmer temperatures. These differences indicate that the relative LR constants presented herein should not be applied to appreciably different (i.e., ≲10 °C) thermal environments without correction for these temperature effects, for which accurate data on enthalpies of vaporization of PAHs from petroleum as a function of temperature would be helpful.

The FO weathering model may be used to predict relative PAH concentrations that evaporate into the atmosphere or that dissolve into aqueous solution. From eq 3, the instantaneous rate of decrease of a PAH P_j from weathering petroleum is $-dP_j/dw = k_jP_j$, and is proportional with the instantaneous increase in the concentration of P_j in the receiving phase. This ensures that PAH concentrations of the receiving phase are correlated with PAH concentrations initially present in the petroleum. Predicted correlation coefficients for aqueous PAH concentrations dissolved from initially unweathered EVO and PAH concentrations initially present in unweathered EVO (proportional to π_j) exceed 0.9, and are consistent with correlation coefficients based on dissolved PAH concentrations measured in seawater 1–2 weeks following the EVOS (*25*). The high correlation is because variability

among π_j is substantially greater than variability among the k_j, and the most rapidly dissolving (or evaporating) PAH tend also to be the most abundant initially present in EVO.

The Weathering Model and PAH Source Identification. Our weathering model is related to currently accepted protocols for oil-spill source identification based on PAH analysis (*26*, *27*). The protocols currently adopted in the United States and in Europe compare normalized PAH results for samples and suspected sources, and patterns that match within constraints imposed by analytical precision and by weathering effects are accepted as evidence implicating the suspected source. The constraints imposed by weathering effects include decreasing trends in pattern discrepancies with increasing PAH boiling points (*27*) or with increasing alkyl-substitution within homologous PAH series (*26*). Those normalized PAHs identified as unaffected by weathering can be included in multivariate statistical comparisons with corresponding results from suspected oil-spill sources to evaluate whether discrepancies that remain among these analytes can be ascribed to analytical precision. The identification procedure thus employs two criteria: patterns of normalized PAH that match within the constraints of analytical precision for analytes that are not affected by weathering, and patterns of PAH weathering losses that conform with specified trends.

Our weathering model may be regarded as an alternative formulation of the above protocols. It provides an explicit mathematical specification of the PAH weathering-loss trends, and the PAHs included in the matching procedure are extended to weathered PAH. The weathering model reduces to a simple comparison of relative PAH concentrations in a sample i and in a suspected source oil as w_i approaches zero. As w_i increases, progressively more PAH are significantly affected by weathering, depending on comparison of the product

k_j w_i and analytical precision for the jth PAH expressed as a coefficient of variation. The European protocol (*27*) makes greater use of the information produced by the chemical analysis, because all the resolvable isomer peaks are considered individually, in contrast to summation of isomer peaks for each homologue reported, as was done here. Although this is a substantial advantage of the European protocol, our model could be adapted to such protocols as a possible refinement.

An advantage unique to the weathering model is its provision for more precise definition of weathering. The weathering parameter w_i defines weathering by indexing the relative abundance of a set of PAH with known LR constants, so that comparisons between samples can be unambiguously controlled for weathering, which leads to a distribution for MSE_i that is independent of weathering state. Given a distribution for MSE_i derived from laboratory observations, source identifications can be evaluated by estimates of the probability of committing type I error. In our model, a type I error is an erroneous rejection of the null hypothesis that the pattern of PAH in a sample is consistent with EVO.

By quantifying the weathering state, w_i also provides an index of the potential toxicity remaining in the oil of a sample. Lower values of w_i indicate progressively greater relative abundances of the PAHs that are most readily lost to the environment, and PAHs are the most toxic components (in absolute terms) of petroleum (*28*). The value of w_i is thus inversely related to the toxic burden remaining in an oil sample. In this regard, w_i is an especially appropriate parameter for bioremediation studies, where the objective is to find biological conditions that accelerate dw_i/dt.

Although we used chrysenes to determine the parameter c_i of unweathered petroleum originally introduced into the *i*th sample, other temporally-invariant constituents of petroleum could also be used. The low estimates of k_j we obtained for C3- and C4-phenanthrenes indicate that these could have been included in eq 4 to estimate c_i, but were not because the chrysenes were chosen *a priori* for this purpose. Alternatively, other persistent constituents such the alicyclic hopanes or stearanes could be used.

Successful application of our weathering model requires careful consideration of the interaction among (a) the PAH set selected for inclusion in the model, (b) the detection limit definition chosen, and (c) the effects of these choices on the scaling of w_i. Detection limit stipulation is critical because the value calculated for the fit of the model to the data (i.e., MSE_i) will increase dramatically if contributions from analytes well below detection limits are included, since the discrepancy between observed and predicted analyte concentrations may be orders of magnitude at concentrations sufficiently below detection limits. Thus, once weathering proceeds to the point where concentrations of one or more of the analytes in the weathering model are below detection limits, application of the model may be compromised. The choice of analytes included in the model, together with the detection limits used, therefore determine the range of weathering states covered by the model. The model may consequently fail to apply to samples that contain very high concentrations of petroleum if it is very weathered, because the dilution necessary for valid analysis of the most abundant analytes included in the model may cause the most weathered analytes to fall below the detection limits applied.

The choice of analytes included in the model also affects the ability of the model to distinguish among initial stages of petroleum weathering, because most of the information regarding w_i is contained in the measurements of the most rapidly lost analytes. Consequently, a model that includes rapidly lost analytes (e.g., naphthalene) will distinguish among earlier weathering stages better one that does not, but the latter will be applicable over a broader range of weathering states, because its constituent analytes are more persistent. Also, the weathering states that correspond to a particular value of w_i will not be the same for these two models, owing to the normalization condition $\sum k_j^2 = 1$. This condition causes the results for w_i to be a function of the analytes included in the model, so w_is based on different analyte sets cannot be directly compared.

The ratio of surface area to volume of petroleum in a sample is an important factor affecting the relation of w_i and time. As demonstrated in Figure 2, dw_i/dt decreases as the film thickness of petroleum applied to the gravel substrate increases; this behavior is consistent with the RLS for PAH-loss from petroleum discussed above. This implies that a variety of weathering states may be observed shortly following an oil spill, depending on the surface area to volume ratio of the petroleum sampled. Also, relatively unweathered petroleum may persist for prolonged periods in the field if the surface area exposed to wind or water currents is small relative to the petroleum volume associated with the matrix sampled, hence the weak correlation of w_i and time (Figure 7).

Assessment of First-Order Weathering Kinetics for Petroleum Spilled from the T/V *Exxon Valdez*. The applicability of the laboratory-derived weathering model to field results from an oil spill may be assessed by (a) comparing the error distributions derived from

applying the weathering model to laboratory results *versus* field results, and (b) comparing the geographic distribution of samples identified as contaminated by EVO with the geographic boundaries of the area contaminated by the EVOS.

The error distributions that result from applying the weathering model to field samples of mussels and sediments confirm that the dominant weathering processes of this oil spill followed FOLR kinetics. The median MSE_k of 0.25 for EVO-contaminated field mussels means that most of the PAH concentrations predicted by the weathering model fall within −40% to +65% of observed concentrations. The corresponding range for the laboratory weathering experiments is −33% to +49%, which indicates that the weathering model is almost as successful at predicting PAH concentrations in field mussels as it is at predicting PAH concentrations in experimentally-weathered petroleum. The greater disparity between observed and predicted PAHs in mussels compared with experimentally-weathered petroleum is probably due to the combined effects of small interferences from other hydrocarbon sources in the environment or introduced during sample collection or storage, small PAH composition differences between the petroleum used for the weathering experiment and the petroleum spilled, and composition differences induced by the differences in the thermal histories of the spilled petroleum and the petroleum used for the weathering experiments. These effects may collectively be regarded as small perturbations compared with the much larger PAH concentration changes that result from FO weathering processes.

The similarly leptokurtic distributions of the MSE_ks of mussels and the MSE_is of experimentally-weathered petroleum, together with the fact that the distribution for mussels includes nearly all the mussels that meet the MDL requirements of the weathering model,

corroborates the similarity of the underlying weathering processes for EVO in mussels in the field and on gravel in the weathering experiments. The 17 mussels that were modeled as not consistent with EVO at the 1% type I error rate is probably the result of truncation of the MSE_i distribution after the median value is increased to 0.25. The MSE_k distribution based on all the mussels that meet the weathering model requirements is, therefore, generally consistent with expectations based on the weathering model, with small allowance for environmental perturbations.

The similarity of the error distributions derived from mussels and from the more contaminated sediments indicates that the weathering model applies equally well for these matrixes. The median MSE_k for mussels is almost identical with the median for sediments identified as EVO-contaminated at total PAH concentrations >750 ng/g, and both distributions are similarly leptokurtic. Thus, with few exceptions, sediment samples that are sufficiently contaminated by EVO that other hydrocarbon sources are negligible in comparison, display patterns of relative PAH concentrations consistent with FOLR kinetics. That the same model produces similar results for such disparate environmental matrixes further validates the kinetics of the underlying weathering processes assumed by the model. Following the EVOS, PAH losses due to weathering followed FOLR kinetics regardless of the great variability of environmental conditions among intertidal mussels, intertidal sediments, subtidal sediments, and EVO-contaminated sediments in transport from the intertidal to the subtidal, at geographic locations of very different aspects. This generality derives from the simple notion that the rate of PAH loss from petroleum is determined by the energy required for PAH molecules to escape from petroleum.

The effects of PAH interferences on the weathering model can be assessed by constructing hypothetical mixtures of PAHs from EVO and alternative sources, and calculating the increase in MSE_i that results from application of the EVO-weathering model. For example, eq 6 may be applied to PAH of a synthetic sample for which a proportion (1 - q) of the total PAH is from EVO and the remaining proportion q is from an alternative PAH source (such as that apparent in the intertidal sediments at Constantine Harbor). The change of the MSE_i as q increases may be used to assess the sensitivity of the weathering model to interference from the alternative PAH source. This procedure may be bootstrapped at fixed values of q to generate a distribution of $MSE_i^{(q)}$ from mixtures of Constantine Harbor and experimentally-weathered EVO samples, analogous to the generation of the distribution for MSE_i from experimentally-weathered EVO samples. The median value of the $MSE_i^{(q)}$ distribution is comparable with the median of the MSE_i distribution of the weathering model for $q \lesssim 0.2$, which indicates that the weathering model is not sensitive to mixtures that contain as much as 20% of the total PAH from the natural source (Figure 9). Similarly, the median value of the $MSE_i^{(q)}{}'$ distribution is comparable with the median of the MSE_i' distribution of the natural source model for $q \lesssim 0.05$ which indicates that the natural source model is not sensitive to mixtures that contain as much as 5% of the total PAH from EVO.

We conclude from these exercises on synthetic mixtures of PAH sources that the models we have presented are most validly applied to samples that contain PAH from a single predominant source. This is the usual case for more heavily contaminated samples collected during catastrophic events such as major oil spills, where the PAH contribution from the catastrophic source predominates. It is also the usual case for pristine environments that

contain PAH from a single natural source, or from multiple sources but in contributions of constant proportions (provided no weathering occurs). However, the validity of these models is compromised as PAH contributions from a suspected predominant source approach contributions from alternative sources, because of the difficulty in distinguishing larger MSE_k values that result from mixtures, and larger MSE_k values that result from stochasticity. Stochastic consequences may be important in this context because of the relatively low precision of the underlying analytical measurements.

The geographic distribution of the samples we identified as contaminated by EVO is generally consistent with the trajectory of the spilled petroleum, indicating an absence of spurious identifications generated by the weathering model. This is supported by the low frequencies of mussel or sediment samples identified as EVO-contaminated that were collected outside the trajectory of the spilled petroleum, or that were collected at sediment epibenthic depths below about 20 m. Also, no mussel or sediment sample collected just before landfall of the spilled EVO was identified as EVO-contaminated.

The EVOS provided a unique opportunity to assess the weathering model because the most heavily contaminated compartments of the affected area were among those least affected by PAH inputs from alternative sources before the spill. Before the EVOS, the PAHs characteristic of EVO were rarely detected in mussels outside Port Valdez in PWS (*12*, *13*). Sediment PAH concentrations derived from natural sources decrease with progressively shallower epibenthic depths, so that total PAH concentrations from these sources rarely exceed 100 ng/g in intertidal sediments and 200 ng/g in subtidal sediments to 20 m depth (*14*, *16*). Weathered EVO was most prevalent in these two environmental compartments: mussels, and

sediments at less than 20 m epibenthic depth. As a result, PAH concentration patterns characteristic of weathered EVO are ubiquitous in mussels that contain sufficient PAH to be evaluated by the weathering model, and common in sediments that have total PAH concentrations exceeding about 750 ng/g. Note that a subtidal sediment sample that contains 750 ng/g total PAH with 200 ng/g from natural sources and 550 ng/g from EVO would most likely be classified as EVO by the weathering model procedure, based on the sensitivity of the model to mixtures from these sources discussed above. Interference from natural sources on the EVO identification procedure we present herein is, therefore, probably negligible in mussels, and also in sediments that contain more than about 750 ng/g total PAH.

The absence of PAH from natural sediment PAH sources in mussels, together with the observation that the PAH pattern that characterizes these sources does not weather, places strong constraints on the nature of these sources. These sources have been identified with natural petroleum seeps along the southern coast of Alaska at Katalla and elsewhere (*16*), but this identification is not obviously consistent with the absence of weathering and with the absence of these PAH in mussels, which implies that these PAH are sequestered in such a way that biological availability is precluded. The absence of weathering and of biological availability is most clearly evident at Constantine Harbor, where concentrations of PAH most susceptible to weathering have not changed in intertidal sediments during a 15-year monitoring period, and were rarely detected in adjacent mussels simultaneously collected. The pattern of PAH concentrations of unburned coal is difficult to distinguish from petroleum in sediments (*29*), and because PAH sequestered in microscopic coal particles is consistent with the absence of weathering and with the absence of these PAH in mussels, coal has been

suggested as an alternative source (*13*, *14*). Coal as a possible source has been dismissed based on the apparent absence of inventoried coal deposits in Alaska [(*30*); although cited in (*31*), this reference does not appear to address this issue] to account for the characteristic PAH pattern that is observed in submarine sediments east of Katalla (*31*). However, absence of proof is not equivalent to proof of absence, and undiscovered coal deposits in Alaska (or discovered deposits in the Alsek river drainage of Canada) remain plausible sources, so this dismissal is premature. Conversely, before a petroleum seep source is accepted, the absence of weathering and of bioavailability must be explained, because these are not consistent with the environmental behavior of petroleum.

Acknowledgments

The research described in this paper was supported by the *Exxon Valdez* Oil Spill Trustee Council. However, 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) Wang, Z.; Fingas, M.; Sergy, G. *Environ. Sci. Technol.* **1994**, *28*, 1733–1746.

(2) Killeen, T. J.; Eastwood, D.; Hendrick, M. S. *Talanta* **1981**, *28*, 1–6.

(3) Venosa, A. D.; Suidan, M. T.; Wrenn, B. A.; Strohmeier, K. L.; Haines, J. R.; Eberhart, B. L.; King, D.; Holder, E. *Environ. Sci. Technol.* **1996**, *30*, 1764–1775.

(4) Singer, M. E.; Finnerty, W. R. In *Petroleum Microbiology*; Atlas, R. M., Ed.; Macmillan publishing Co.: New York, 1984.

(5) Roques, D. E.; Overton, E. B.; Henry, C. B. *J. Environ. Qual.* **1994**, *23*, 851–855.

(6) Michel, J.; Hayes, M. O. In *Proceedings of the 1993 International Oil Spill Conference*; American Petroleum Institute: Washington, DC, 1993.

(7) Kennicutt, M. C., II. *Oil Chem. Pollut.* **1988**, *4*, 89–112.

(8) Boehm, P. D.; Steinhauer, M. S.; Green, D. R.; Fowler, B.; Humphrey, B.; Fiest, D. L.; Cretney, W. J. *Arctic* **1987** *40* (Suppl. 1), 133–148.

(9) Humphrey, B.; Boehm, P. D.; Hamilton, M. C.; Norstrom, R. J. *Arctic* **1987** *40* (Suppl. 1), 149–161.

(10) Majewski, J.; O'Brien, J. *Environmental Letters* **1974**, *7* (2), 145–161.

(11) Hansen, H. P. *Rapp. P.-v. Reun. Cons. int. Explor. Mer* **1977**, *171*, 101–106.

(12) Karinen, J. F.; Babcock, M. M.; Brown, D. W.; MacLeod, W. D., Jr.; Ramos, L. S.; Short, J. W. *Hydrocarbons in Intertidal Sediments and Mussels from Prince william Sound, Alaska, 1977–1980: Characterization and Probable Sources;* U.S. Department of Commerce, Auke Bay Lab: Juneau, AK, 1993; NOAA Tech. Memo. NMFS-AFSC-9.

(13) Short, J. W.; Babcock, M. M. *Am. Fish. Soc. Symp.* **1996**, *18*, 149–166.

(14) O'Clair, C. E.; Short, J. W.; Rice, S. D. *Am. Fish. Soc. Symp.* **1996**, *18*, 61–93.

(15) Short, J. W.; Sale, D. M.; Gibeaut, J. C. *Am. Fish. Soc. Symp.* **1996**, *18*, 40–60.

(16) Page, D. S.; Boehm, P. D.; Douglas, G. S.; Bence, A. E.; In *Third ASTM Symposium on Environmental Toxicology and Risk Assessment: Aquatic, Plant, and Terrestrial;* April 25–28, 1993, Atlanta, GA; American Society for Testing and Materials: Philadelphia, PA, 1993.

(17) Short, J. W.; Jackson, T. J.; Larsen, M. L.; Wade, T. L. *Am. Fish. Soc. Symp.* **1996**, *18*, 140–148.

(18) Marty G. D.; Short, J. W.; Dambach D. M.; Willits, N. H.; Heintz, R. A.; Rice, S. D.; Stegeman, J. J.; Hinton, D. E. Submitted, *Can. J. Zool.*

(19) Glaser, J. A.; Forest, D. L.; McKee, G. D.; Quave, S. A.; Budde, W. L. *Environ. Sci. Technol.* **1981**, *15*, 1426–1435.

(20) Wolfe, D. A.; Hameedi, M. J.; 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) Niebauer, H. J.; Royer, T. C.; Weingartner, T. J. *J. Geophys. Res.* **1994**, *99*, 14,113–14,126.

(22) Wold, S.; Sjöström, M. In *Chemometrics: Theory and Application*; Kowalski, B. R., Ed.; ACS Symposium Series 52; Washington, DC, 1977.

(23) Sale, D. M.; Gibeaut, J. C.; Short, J. W. *Nearshore Transport of Hydrocarbons and Sediments Following the Exxon Valdez Oil Spill, Exxon Valdez Oil Spill State/Federal Natural Resource Damage Assessment Final Report (Subtidal Study Number 3B)*; Alaska Department of Environmental Conservation: Juneau, AK, 1995.

(24) Bence, A. E.; Burns, W. A. 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: Philadelphia, PA, 1995; special technical publication 1219.

(25) Short, J. W.; Harris, P. M. *Am. Fish. Soc. Symp.* **1996**, *18*, 17–28.

(26) ASTM. In *Annual Book of ASTM Standards, Section 11: Water and Technology* **1996**, *11.02* Water (II), 835–845.

(27) NORDTEST. *Nordtest method*; NT Chem 001, Ed. 2; NORDTEST: Espoo, Finland, 1991.

(28) National Research Council. *Oil in the Sea: Inputs, Fates, and Effects*; National Academy Press: Washington, DC, 1985 (p. 372).

(29) Tripp, B. W.; Farrington, J. W.; Teal, J. M. *Mar. Pollut. Bull.* **1981**, *12*, 122–126.

(30) Tysdal, R. G.; Case, J. E. *Geologic Map of the Seward and Blying Sound Quadrangles, Alaska. Miscellaneous Investigations Series Map I-1150*; U.S. Geological Survey: Washington, DC, 1979. [Referred to in (*31*).]

(31) Page, D. S.; Boehm, P. D.; Douglas, G. S.; Bence, A. E.; Burns, W. A.; Mankiewicz, P. J. *Environ. Toxicol. Chem.* **1996**, *15*, 1266–1281.

(32) Pearlman, R. S.; Yalkowsky, S. H.; Banerjee, S. *J. Phys. Chem. Ref. Data* **1984**, *13*, 555–562.

Appendix

Least-square estimation of w_i and k_j

Least square (LS) estimates of w_i and k_j satisfy the following conditions:

$$\frac{\partial}{\partial k_j}\{\text{trace}[(\mathbf{D} - \hat{\mathbf{w}}\hat{\mathbf{k}})'(\mathbf{D} - \hat{\mathbf{w}}\hat{\mathbf{k}})]\} = 0 \qquad \textbf{(A1)}$$

and

$$\frac{\partial}{\partial w_i}\{\text{trace}[(\mathbf{D} - \hat{\mathbf{w}}'\hat{\mathbf{k}})'(\mathbf{D} - \hat{\mathbf{w}}'\hat{\mathbf{k}})]\} = 0 \qquad \textbf{(A2)}$$

where $\hat{\mathbf{w}}'$ is a column vector with elements $\hat{w}_i$ and $\hat{\mathbf{k}}$ is a row vector with elements $\hat{k}_j$. Because the elements of $\mathbf{D}$ are measured constants here, and because trace $\mathbf{D}'\hat{\mathbf{w}}'\hat{\mathbf{k}}$ = trace $\hat{\mathbf{k}}'\hat{\mathbf{w}}\mathbf{D}$, these two conditions lead to

$$2\frac{\partial}{\partial k_j}(\text{trace } \mathbf{D}'\hat{\mathbf{w}}'\hat{\mathbf{k}}) = \frac{\partial}{\partial k_j}(\text{trace } \hat{\mathbf{k}}'\hat{\mathbf{w}}\hat{\mathbf{w}}'\hat{\mathbf{k}}) \qquad \textbf{(A3)}$$

and

$$2\frac{\partial}{\partial w_i}(\text{trace } \mathbf{D}'\hat{\mathbf{w}}'\hat{\mathbf{k}}) = \frac{\partial}{\partial w_i}(\text{trace } \hat{\mathbf{k}}'\hat{\mathbf{w}}\hat{\mathbf{w}}'\hat{\mathbf{k}}) \qquad \textbf{(A4)}$$

Noting that

$$\text{trace } \mathbf{D}'\hat{\mathbf{w}}'\hat{\mathbf{k}} = \sum_{j=1}^{J}\left(\hat{k}_j\sum_{i=1}^{I} d_{ij}\hat{w}_i\right) \qquad \textbf{(A5)}$$

and

$$\text{trace } \hat{\mathbf{k}}'\hat{\mathbf{w}}\hat{\mathbf{w}}'\hat{\mathbf{k}} = \left(\sum_{i=1}^{I}\hat{w}_i^2\right)\left(\sum_{j=1}^{J}\hat{k}_j^2\right) \qquad \textbf{(A6)}$$

the differentiations indicated in eqs A3 and A4 produce the following equations for the LS estimates of the parameters $\hat{k}_j$ and $\hat{w}_i$, respectively:

$$\hat{k}_j = \frac{\sum_{i=1}^{I} \hat{w}_i d_{ij}}{\sum_{i=1}^{I} \hat{w}_i^2} \qquad \textbf{(A7)}$$

and

$$\hat{w}_i = \frac{\sum_{j=1}^{J} \hat{k}_j d_{ij}}{\sum_{j=1}^{J} \hat{k}_j^2} \qquad \textbf{(A8)}$$

Equations A7 and A8 imply that $\hat{\mathbf{k}}'$ is an eigenvector of $\mathbf{D}'\mathbf{D}$ as follows: given

$$e_{ij} = d_{ij} - \hat{w}_i \hat{k}_j \; ; \qquad \sum_{j=1}^{J} e_{ij}\hat{k}_j = \sum_{j=1}^{J} d_{ij}\hat{k}_j - \hat{w}_i \sum_{j=1}^{J} \hat{k}_j^2 = 0$$

by eq A7, so

$$\hat{\mathbf{E}}\hat{\mathbf{k}}' = \mathbf{0}^{I\times 1} \qquad \textbf{(A9)}$$

Also,

$$\sum_{i=1}^{I} \hat{e}_{ij}\hat{w}_i = \sum_{i=1}^{I} \left(d_{ij} - \hat{w}_i \hat{k}_j\right) \hat{w}_i = \sum_{i=1}^{I} d_{ij}\hat{w}_i - \hat{k}_j \sum_{i=1}^{I} \hat{w}_i^2 = 0 \qquad \textbf{(A10)}$$

by eq. 17, so

$$\hat{\mathbf{w}}\hat{\mathbf{E}} = \mathbf{0}^{1\times J} \qquad \textbf{(A11)}$$

Now

$$\mathbf{D}'\mathbf{D} = (\hat{\mathbf{w}}'\hat{\mathbf{k}} + \hat{\mathbf{E}})'(\hat{\mathbf{w}}'\hat{\mathbf{k}} + \hat{\mathbf{E}}) = \hat{\mathbf{k}}'\hat{\mathbf{w}}\hat{\mathbf{w}}'\hat{\mathbf{k}} + \hat{\mathbf{k}}'\hat{\mathbf{w}}\hat{\mathbf{E}} + \hat{\mathbf{E}}'\hat{\mathbf{w}}'\hat{\mathbf{k}} + \hat{\mathbf{E}}'\hat{\mathbf{E}}$$

Multiplication of **D'D** on the right by $\hat{\mathbf{k}}'$, and using eqs A9 and A11 gives

$$\mathbf{D}'\mathbf{D}\,\hat{\mathbf{k}}' = \hat{\mathbf{k}}'\sum_{i=1}^{I}\hat{w}_i^2 \qquad \textbf{(A12)}$$

if $\sum k_j^2 = 1$, showing that $\hat{\mathbf{k}}'$ is an eigenvector of **D'D**, associated with the following eigenvalue:

$$\lambda_1 = \sum_{i=1}^{I}\hat{w}_i^2 \qquad \textbf{(A13)}$$

This eigenvalue λ_1 is the dominant eigenvalue of **D'D**, because the remaining eigenvectors of **D'D** form a basis for the error space of $\hat{\mathbf{E}}$, the elements of which are minimized by the LS procedure. The constraint $\sum k_j^2 = 1$ means that the k_j are all relative to an arbitrary scaling factor.

TABLE 1.

PAHs determined in environmental samples collected for the EVOS, and PAH proportions in petroleum spilled from the T/V *Exxon Valdez* in Prince William Sound on 24 March 1989. The 14 polynuclear aromatic hydrocarbons that persisted above method detection limits (MDL; see text) during the 6-month weathering experiments are indicated by *, and the corresponding proportions by weight ($\times 10^3$) are taken as the π_j for the weathering model. Also listed are PAH abbreviations used in the figures of this report, and coefficients of variation for the analysis of these PAH in reference samples at the Auke Bay Laboratory (N = 102). ND = not determined, concentration below MDL in reference sample.

PAHs	Abbreviation	$\pi_j(\times 10^3)$	Coefficient of Variation (%)
Naphthalene		0.724	7.41
2-Methylnaphthalene		1.33	6.62
1-Methylnaphthalene		1.02	6.17
Biphenyl		0.183	12.0

C-2 Naphthalenes		3.15	19.4
Acenaphthylene		0.0139	8.99
Acenaphthene		0.0174	20.0
*C-3 Naphthalenes	C3naph	2.35	11.3
*C-4 Naphthalenes	C4naph	0.598	23.8
Fluorene		0.0911	17.7
C-1 Fluorenes		0.225	27.8
*C-2 Fluorenes	C2fluor	0.191	34.2
*C-3 Fluorenes	C3fluor	0.151	50.7
Dibenzothiophene		0.195	16.3
*C-1 Dibenzothiophenes	C1dithio	0.417	16.7
*C-2 Dibenzothiophenes	C2dithio	0.570	19.1
*C-3 Dibenzothiophenes	C3dithio	0.481	35.4
Phenanthrene		0.255	13.4
Anthracene		<0.001	19.1
*C-1 Phenanthrene/Anthracenes	C1phenan	0.755	20.1
*C-2 Phenanthrene/Anthracenes	C2phenan	0.892	12.0
*C-3 Phenanthrene/Anthracenes	C3phenan	0.558	15.5
*C-4 Phenanthrene/Anthracenes	C4phenan	0.166	41.1
Fluoranthene		0.00909	18.5
Pyrene		0.0147	12.6
C-1 Fluoranthene/Pyrenes		0.0716	16.8

Benz-a-anthracene		<0.001	20.0
*Chrysene	Chrysene	0.0492	20.5
*C-1 Chrysenes	C1chrys	0.0802	23.3
*C-2 Chrysenes	C2chrys	0.106	38.4
C-3 Chrysenes		0.0362	43.2
C-4 Chrysenes		<0.001	ND
Benzo-(b+k)-fluoranthene		0.00644	19.8
Benzo-e-pyrene		0.0119	21.6
Benzo-a-pyrene		<0.001	19.7
Perylene		<0.001	19.7
Indeno[1,2,3-cd]pyrene		<0.001	21.2
Dibenzo-a,h-anthracene		<0.001	43.0
Benzo-ghi-perylene		<0.001	23.9

TABLE 2.

Loss-rate (LR) constants $\hat{k}$ derived from principal component analysis of PAH data for the petroleum weathering experiment. Parentheses contain the range of the central 95% of results from bootstrap iterations of LR constant estimates (see text).

PAH	$\hat{k}$
C-3 Naphthalenes	0.659 (0.653, 0.706)
C-4 Naphthalenes	0.148 (0.040, 0.215)
C-2 Fluorenes	0.118 (0.003, 0.188)
C-3 Fluorenes	0.082 (0.013, 0.147)
C-1 Dibenzothiophenes	0.433 (0.392, 0.462)
C-2 Dibenzothiophenes	0.188 (0.144, 0.258)
C-3 Dibenzothiophenes	0.056 (0.019, 0.119)
C-1 Phenanthrene/Anthracenes	0.512 (0.456, 0.585)
C-2 Phenanthrene/Anthracenes	0.126 (0.086, 0.181)
C-3 Phenanthrene/Anthracenes	-0.027 (-0.068, 0.011)
C-4 Phenanthrene/Anthracenes	-0.024 (-0.055, 0.005)
Chrysene	0.041 (0.015, 0.064)

C-1 Chrysenes	-0.051 (-0.068, -0.034)
C-2 Chrysenes	0.036 (0.007, 0.076)

Figure Legends

FIGURE 1. Map of the northern Gulf of Alaska showing the area affected by the Exxon Valdez oil spill of 24 March 1989 (shaded region). Arrows indicate the path of the Alaska Coastal Current, which flushes Prince William Sound (PWS) and transports sediments from the Copper River and eastward into PWS.

FIGURE 2. Regression relations of weathering parameters (w) and time (days) at four loadings of petroleum on gravel used in the petroleum weathering experiments. Petroleum loadings are expressed as ng total PAH per g gravel.

FIGURE 3. Bootstrapped frequency and cumulative distribution of MSE_i derived from the fit of the bootstrapped iterations of the weathering model to the PAH data from the petroleum weathering experiments. The arrow indicates the MSE_i at the 99th percentile of the cumulative distribution, which is used as a critical value to evaluate the probability that PAH patterns of environmental samples are consistent with weathered EVO.

FIGURE 4. (A) Normalized PAH proportions of unweathered EVO. (B) Predicted and observed normalized PAH proportions of weathered EVO for the case $\hat{w}_i = 3.95$ and $MSE_i = 2.03$, the median of the bootstrap MSE_i distribution. (C) Normalized PAH proportions of sediments from Constantine Harbor, where thin vertical bars indicate the range of the central 95% of results from the bootstrap distribution about the median

indicated by the thick vertical bars. In each case, normalization means that the presented PAH proportions sum to unity.

FIGURE 5. Bootstrapped frequency and cumulative distribution of MSE_i' derived from the fit of the bootstrapped iterations of PAH data from intertidal sediments of Constantine Harbor. The arrow indicates the MSE_i' at the 99th percentile of the cumulative distribution, which is used as a critical value to evaluate the probability that PAH patterns of environmental samples are consistent with the natural PAH source evident at Constantine Harbor.

FIGURE 6. Source classification of PAH in environmental samples as a proportion of samples collected (solid bars) and as a proportion of the sum of the PAH concentrations detected above MDL (shaded bars) for (A) sediments, (B) mussels, and (C) other tissues. The numbers of samples are listed above the solid bars indicating proportions of samples. Sources include EVO = petroleum spilled from the T/V *Exxon Valdez*, Constantine Harbor = the natural sediment PAH source represented by PAH at Constantine Harbor, Neither = other unknown sources (or possibly mixtures of EVO and the natural sediment source), Both = samples that are ambiguously classified, and Not Considered = samples in which one or more of the PAH used in the weathering model are below MDL. For source classification criteria, see text.

FIGURE 7. Weathering parameter $\hat{w}_i$ for EVO-contaminated sediments and mussels versus sample collection time t (in total days) after the EVOS, 1989 to 1995. The linear regression is $\hat{w}_i = 3.557 + 0.000645t$, $r^2 = 0.045$, $P < 0.001$.

FIGURE 8. Arrhenius plot of logarithms of rate-loss constants ($\hat{k}$) from (A) the petroleum weathering experiment and from (B) an independent field experiment (*3*), vs total molecular surface area (TSA) for selected PAH. The selected PAH are identified by abbreviations listed in Table 1, and include the least persistent PAH of the petroleum weathering experiment. Estimates of TSA are presented as nm^2 based on (*32*) for unsubstituted homologues, with 0.20, 0.19, and 0.10 nm^2 added respectively for 1, 2, and each successive carbon of an alkyl substituent [based on average TSA increases due to methyl substitution in (*32*)]. The TSA for dibenzothiophene is estimated as that of fluorene increased by 0.011 nm^2 to account for the longer carbon-sulfur bonds. The TSA is used here as an approximate surrogate measure of vaporization enthalpy. Both sets of rate-loss constants are normalized so that $\Sigma k_j^2 = 1$.

FIGURE 9. Effect of hypothetical mixtures of PAH from EVO and the natural sediment PAH source on the median value of mean square errors (*MSE*) distributions describing the fit of such samples to (A) the EVO weathering model, and (B) the natural sediment PAH source represented by PAH at Constantine Harbor. The abscissa is the proportion $(1 - q)$ of total PAH derived from EVO that is combined with the complementary proportion q derived from the natural sediment source. Random

pairwise combinations according to these proportions of samples from the experimental weathering samples and the Constantine Harbor sediment samples were evaluated by eq 6 & 7 to generate a bootstrapped distribution of the $MSE^{(q)}$, and the median value of these distributions is given as the ordinate.

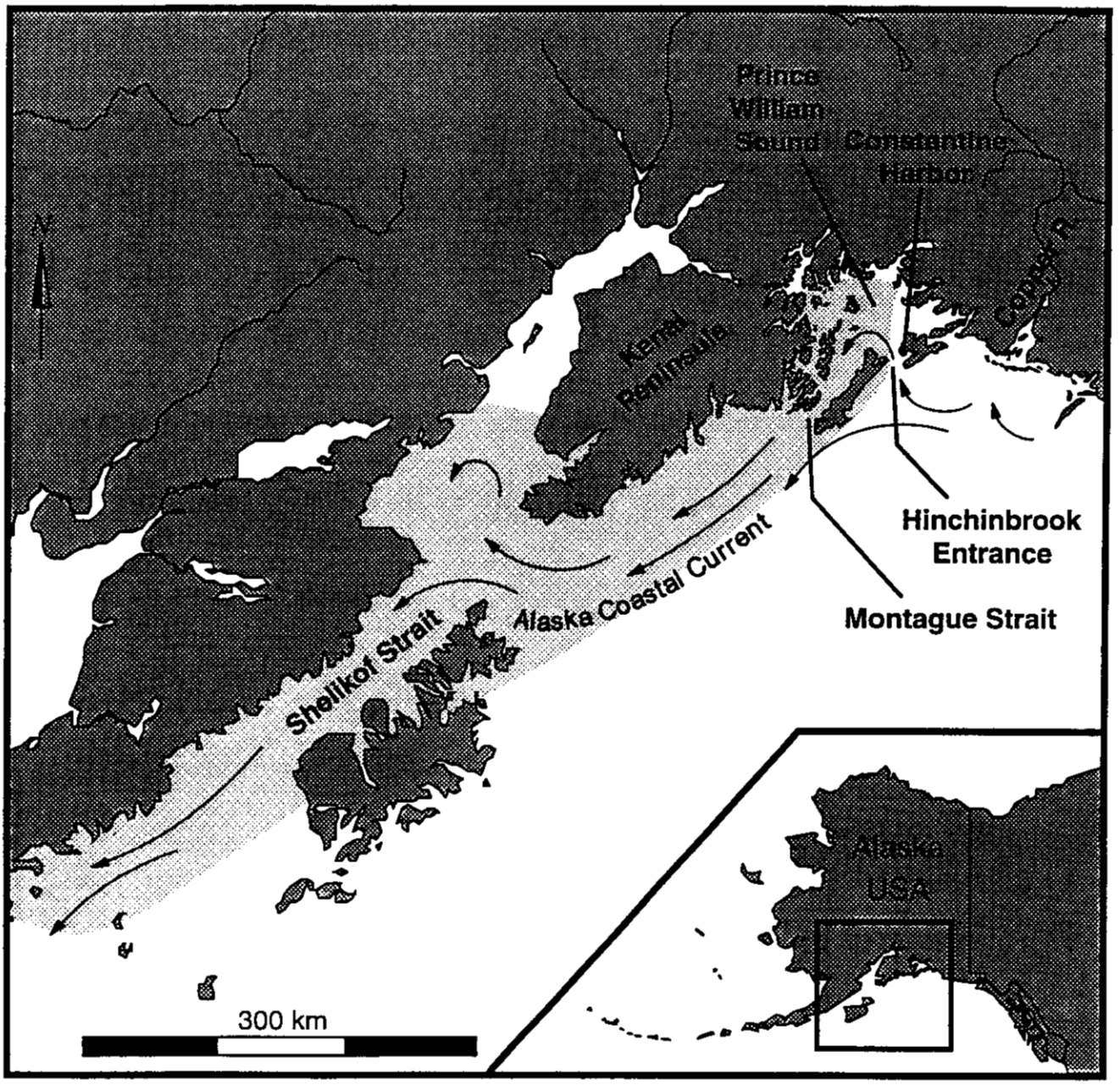

FIGURE 1. The path followed by the spilled *Exxon Valdez* petroleum conformed with the Alaska Coastal Current (ACC).

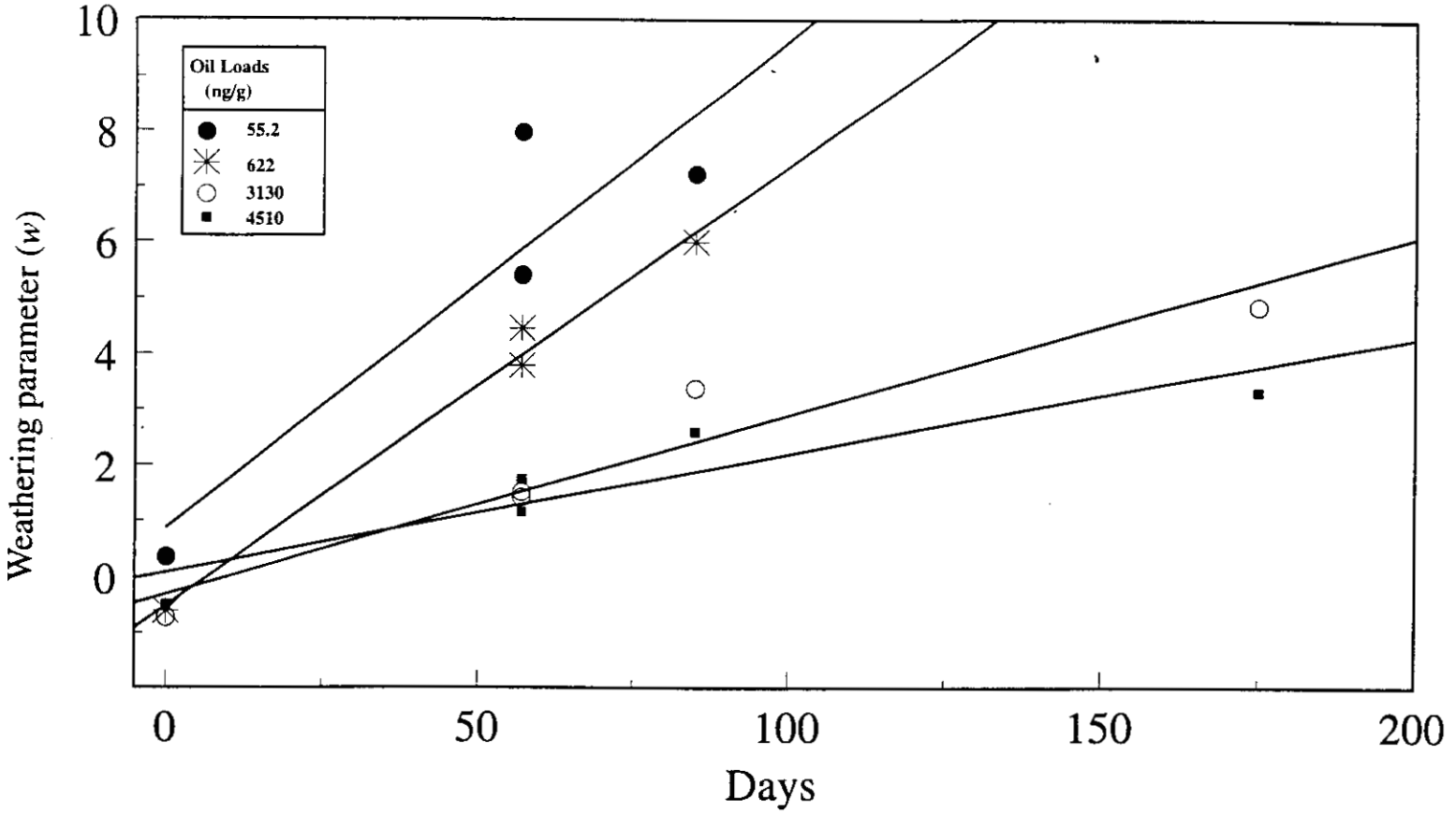

FIGURE 2. Regression relations of weathering parameters (*w*) and time (days) at four loadings of petroleum on gravel used in the petroleum weathering experiments. Petroleum loadings are expressed as ng total PAH per g gravel.

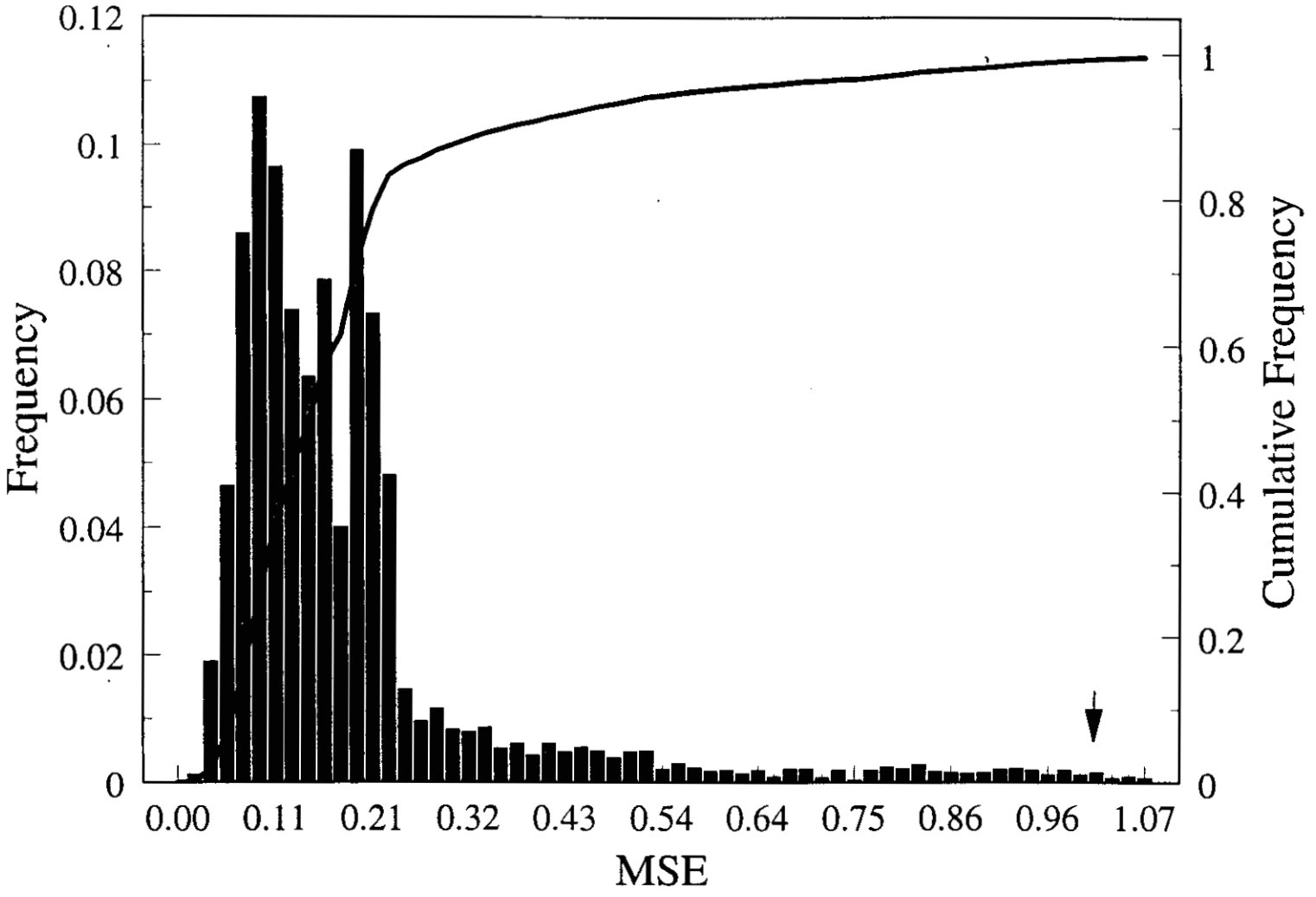

FIGURE 3. Bootstrapped frequency and cumulative distribution of MSEi derived from the fit of the bootstrapped iterations of the weathering model to the PAH data from the petroleum weathering experiments. The arrow indicates the MSEi at the 99th percentile of the cumulative distribution, which is used as a critical value to evaluate the probability that PAH patterns of environmental samples are consistent with weathered EVO.

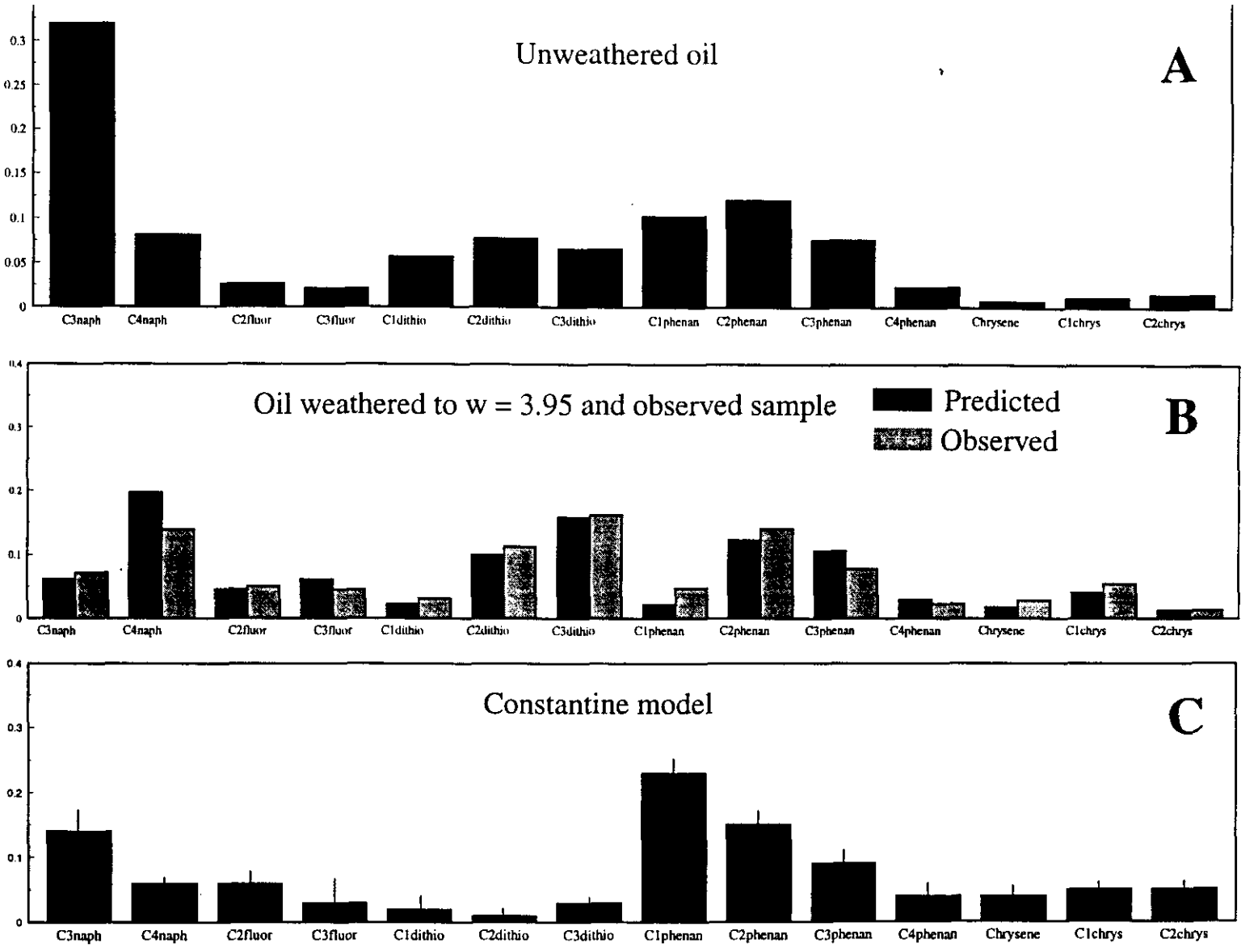

FIGURE 4. (A) Normalized PAH proportions of unweathered EVO. (B) Predicted and observed normalized PAH proportions of weathered EVO for the case i = 3.95 and MSEi = 2.03, the median of the bootstrap MSEi distribution. (C) Normalized PAH proportions of sediments from Constantine Harbor, where thin vertical bars indicate the range of the central 95% of results from the bootstrap distribution about the median indicated by the thick vertical bars. In each case, normalization means that the presented PAH proportions sum to unity.

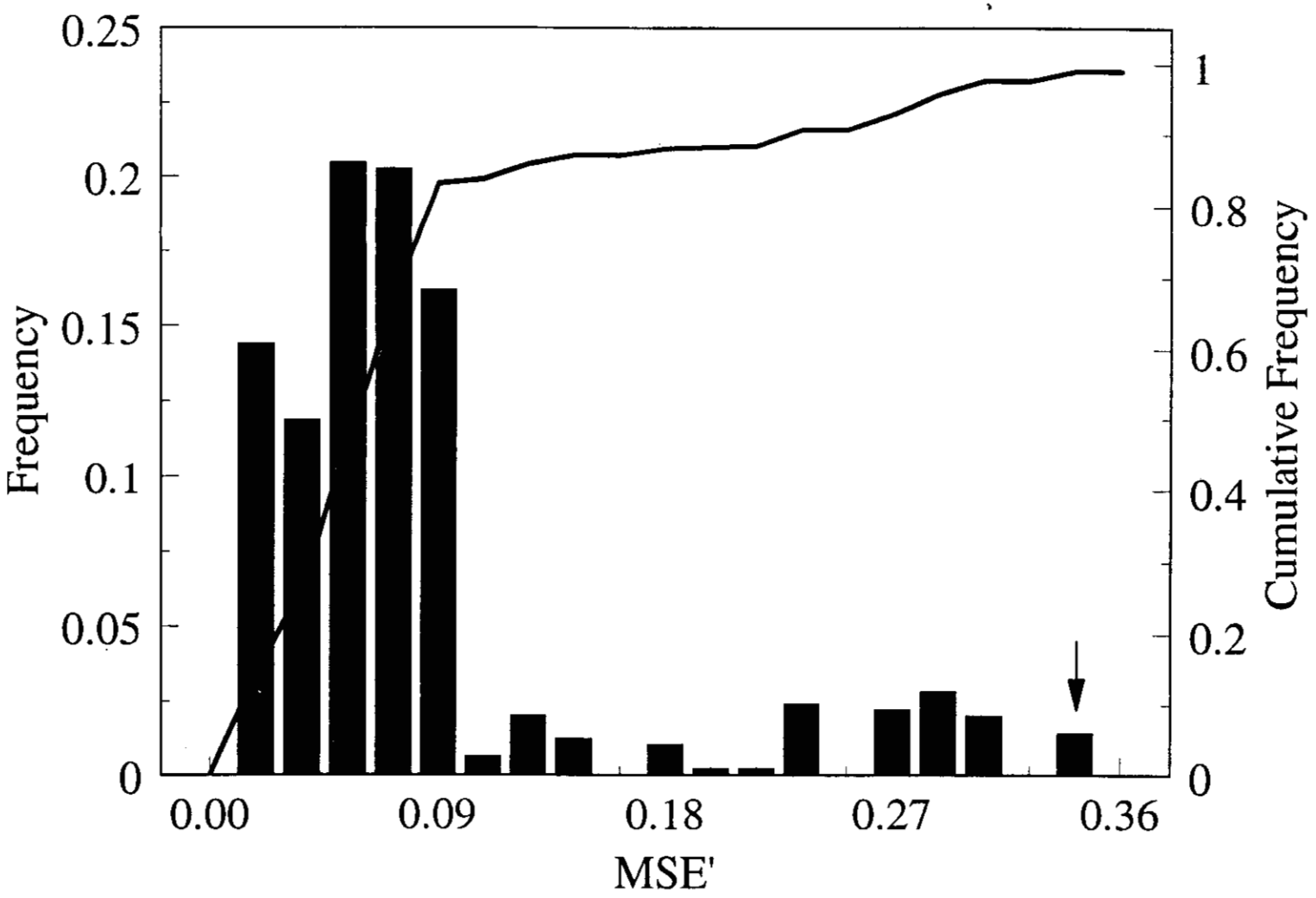

FIGURE 5. Bootstrapped frequency and cumulative distribution of MSEi derived from the fit of the bootstrapped iterations of PAH data from intertidal sediments of Constantine Harbor. The arrow indicates the MSEi at the 99th percentile of the cumulative distribution, which is used as a critical value to evaluate the probability that PAH patterns of environmental samples are consistent with the natural PAH source evident at Constantine Harbor.

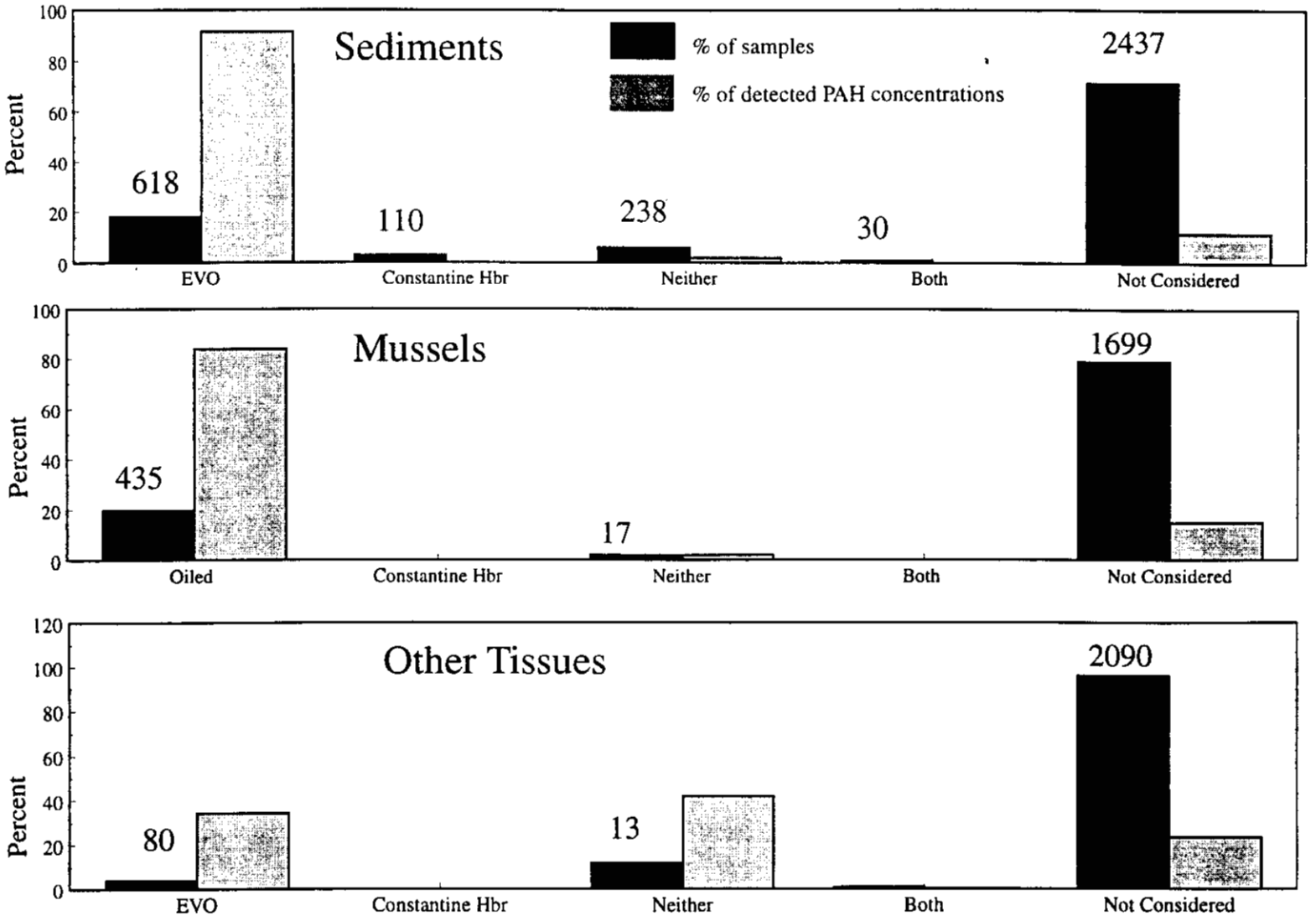

FIGURE 6. Source classification of PAH in environmental samples as a proportion of samples collected (solid bars) and as a proportion of the sum of the PAH concentrations detected above MDL (shaded bars) for (A) sediments, (B) mussels, and (C) other tissues. The numbers of samples are listed above the solid bars indicating proportions of samples. Sources include EVO = petroleum spilled from the T/V Exxon Valdez, Constantine Harbor = the natural sediment PAH source represented by PAH at Constantine Harbor, Neither = other unknown sources (or possibly mixtures of EVO and the natural sediment source), Both = samples that are ambiguously classified, and Not Considered = samples in which one or more of the PAH used in the weathering model are below MDL For source classification criteria, see text.

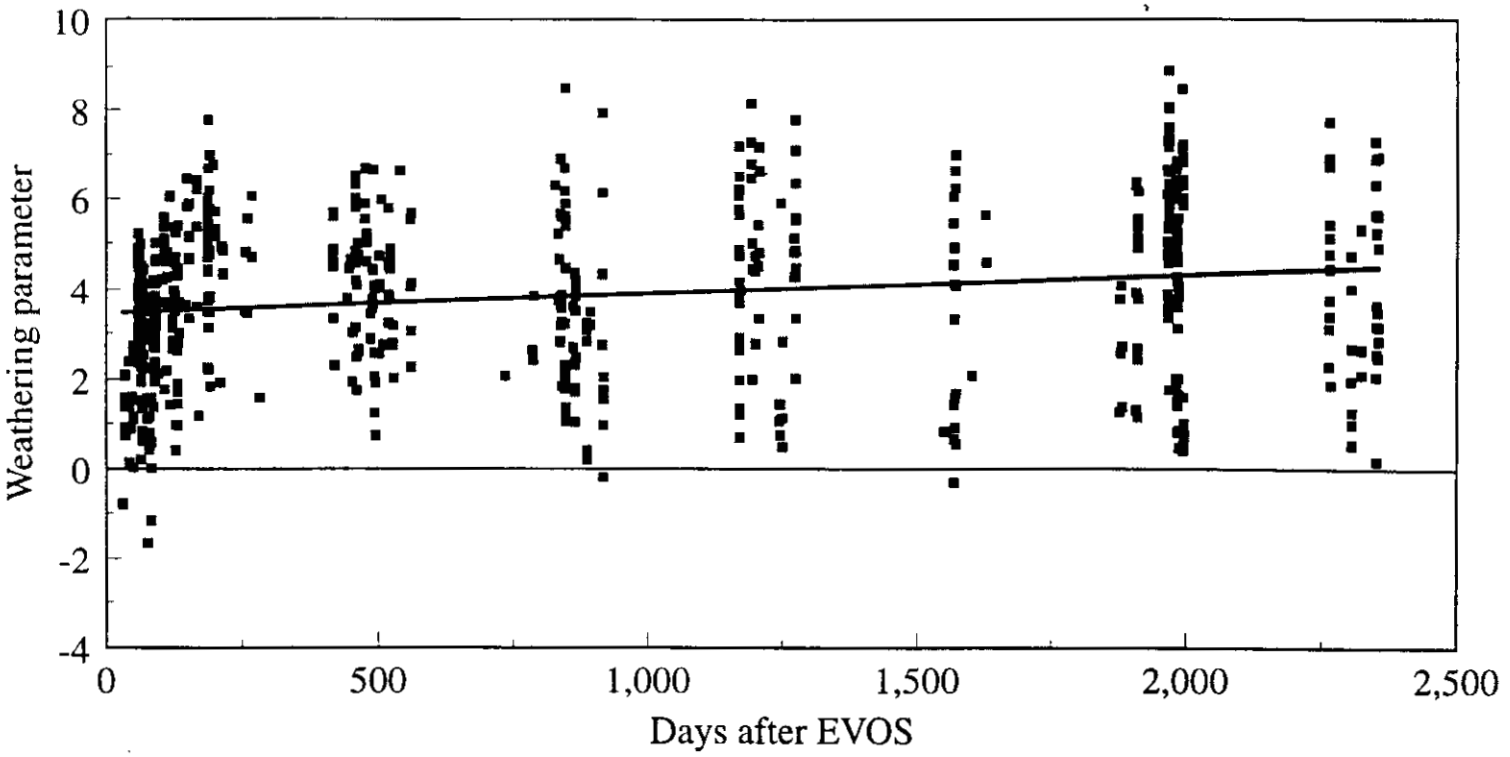

FIGURE 7. Weathering parameter i for EVO-contaminated sediments and mussels versus sample collection time t (in total days) after the EVOS, 1989 to 1995. The linear regression is $i = 3.557 + 0.000645t$, $r2 = 0.045$, $P < 0.001$.

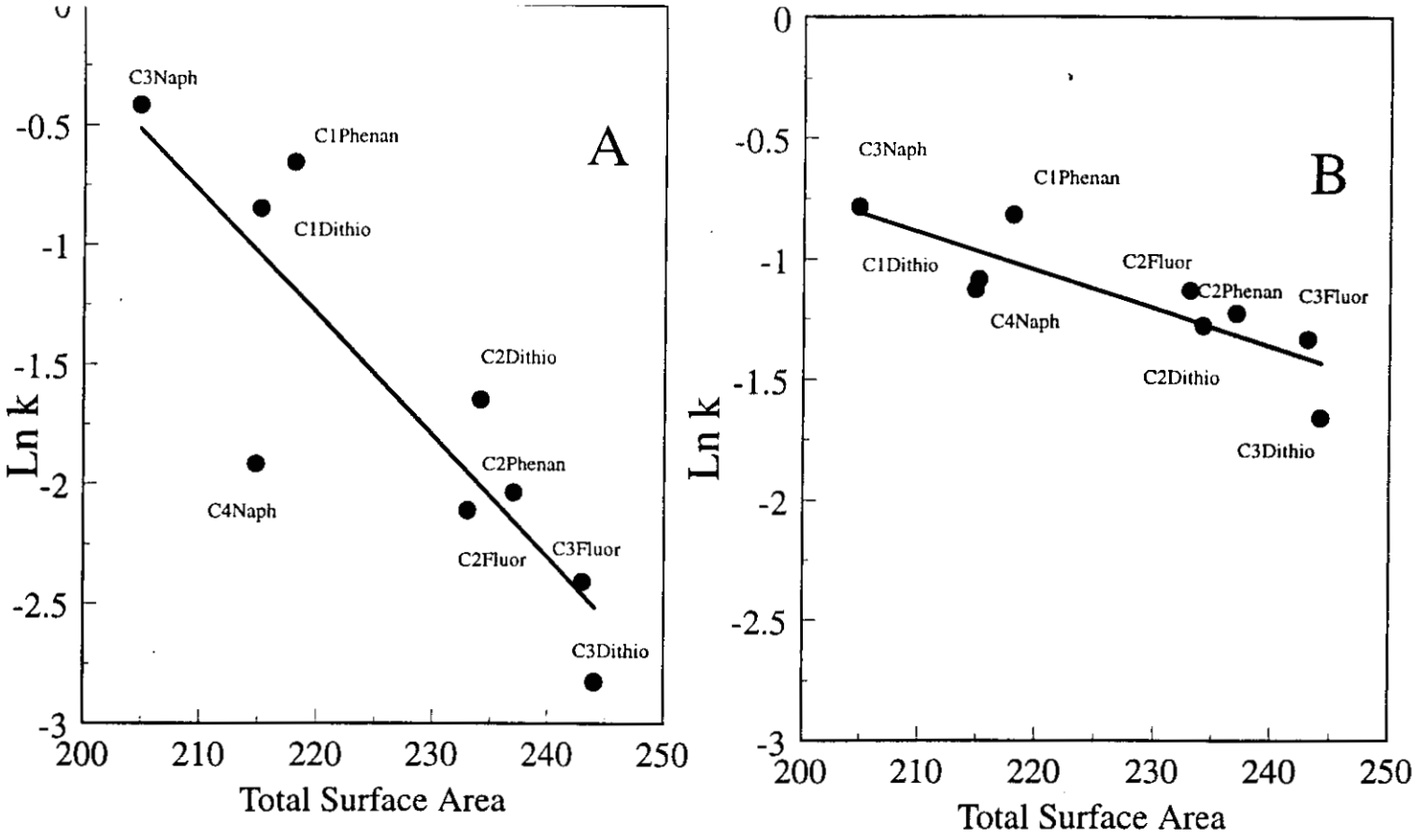

FIGURE 8. Arrhenius plot of logarithms of rate-loss constants () from (A) the petroleum weathering experiment and from (B) an independent field experiment (3), vs total molecular surface area (TSA) for selected PAH. The selected PAH are identified by abbreviations listed in Table 1, and include the least persistent PAH of the petroleum weathering experiment. Estimates of TSA are presented as nm2 based on (32) for unsubstituted homologues, with 0.20, 0.19, and 0.10 nm2 added respectively for 1, 2, and each successive carbon of an alkyl substituent [based on average TSA increases due to methyl substitution in (32)]. The TSA for dibenzothiophene is estimated as that of fluorene increased by 0.011 nm2 to account for the longer carbon-sulfur bonds. The TSA is used here as an approximate surrogate measure of vaporization enthalpy. Both sets of rate-loss constants are normalized so that kj2 = 1.

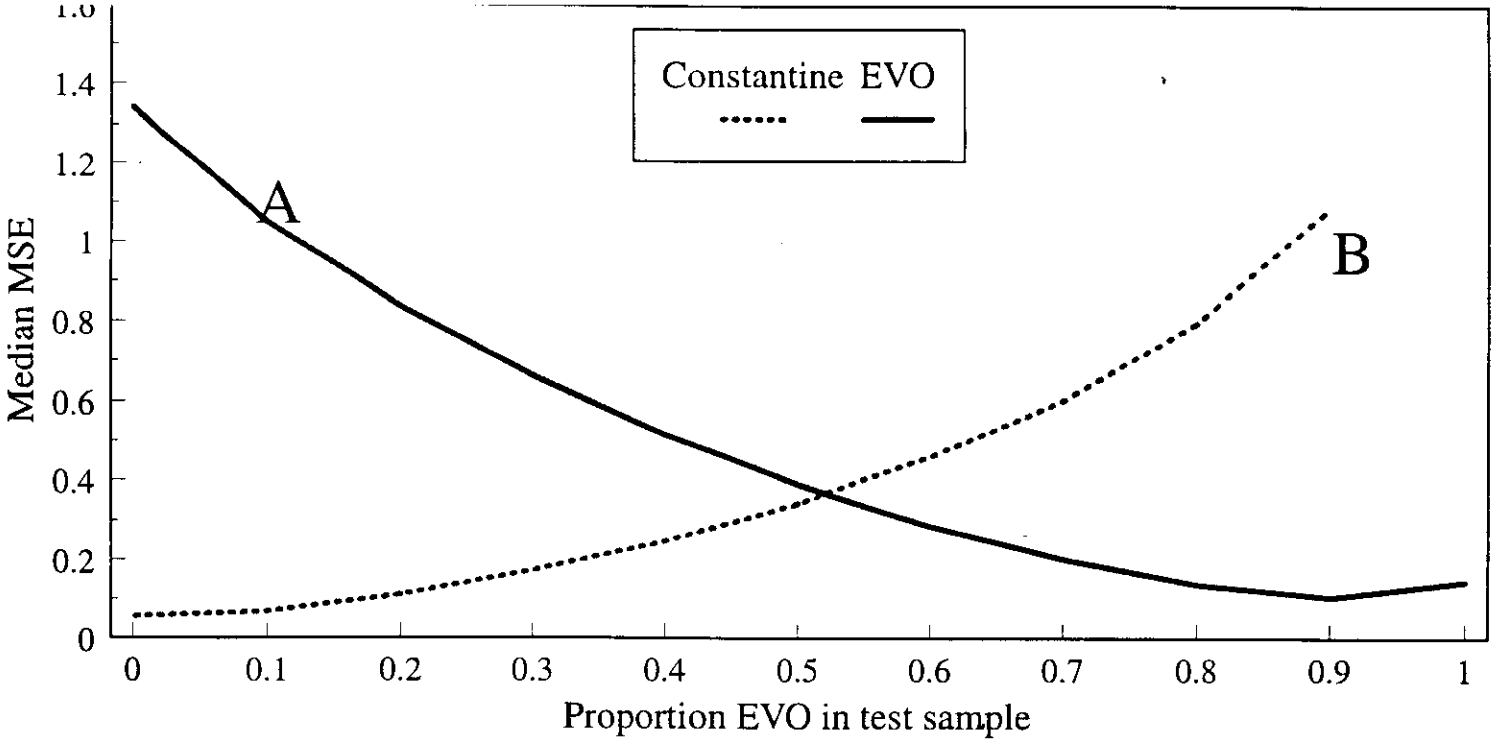

FIGURE 9. Effect of hypothetical mixtures of PAH from EVO and the natural sediment PAH source on the median value of mean square errors (MSE) distributions describing the fit of such samples to (A) the EVO weathering model, and (B) the natural sediment PAH source represented by PAH at Constantine Harbor. The abscissa is the proportion (1 q) of total PAH derived from EVO that is combined with the complementary proportion q derived from the natural sediment source. Random pairwise combinations according to these proportions of samples from the experimental weathering samples and the Constantine Harbor sediment samples were evaluated by eq 6 & 7 to generate a bootstrapped distribution of the MSE(q), and the median value of these distributions is given as the ordinate.

Part III

Descriptive Documentation for Identification of Biased Sediment and Mussel Tissue Samples in the *NRDA* Hydrocarbon Database

Exxon-Valdez Natural Resource Damage Assessment Project

Subtidal 8

DESCRIPTIVE DOCUMENTATION

FOR

IDENTIFICATION OF BIASED SEDIMENT AND MUSSEL TISSUE SAMPLES IN THE *NRDA* HYDROCARBON DATABASE

Prepared by

J. Short and R. Heintz

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

IDENTIFICATION OF BIASED SEDIMENT AND MUSSEL TISSUE SAMPLES IN THE *NRDA* HYDROCARBON DATABASE

The credibility of chemical data derived from an environmental sampling program depends, in part, on the agreement of results among replicate samples, and on confirmation of the general absence of contaminant levels in samples of known uncontaminated sites. Contaminant levels determined by chemical analysis that vary by orders of magnitude among replicate samples, or that occur sporadically among control site samples, may be rightly regarded with skepticism, especially if the outlier samples responsible for these deviations are associated with artificial predictor variables, such as the person who collected the samples, or particular batches of samples analyzed, or the analytical facility, etc. On the other hand, the underlying distribution of contaminants in the sampled environment may be such that levels measured in replicate samples may fail to agree within orders of magnitude in some proportion of the replicated samples. However, in this latter case the outlier samples will be distributed approximately randomly among artificial predictor variables, especially if the data set contains a relatively large number of replicated samples.

The following procedure has been developed to determine whether outlier hydrocarbon data from Exxon-Valdez NRDA samples of sediments or of marine mussel tissue are approximately randomly

distributed among certain artificial predictor variables. The purpose of this procedure is to identify artificial predictor variables that are associated with an improbably large number of outlier samples on the hypothesis of random distribution, so that data from all the samples, whether replicated or not, associated with the identified variable may be used with appropriate caution.

The success of the following procedure depends critically on the relatively large number of samples analyzed and replicated, the relatively large number of chemical analytes simultaneously measured in each sample, and on rigorously consistent definitions of outlier samples. As of October 23, 1992, the Exxon-Valdez NRDA Hydrocarbon Database contained chemical analysis data for 2,698 sediment samples, which includes 1,902 samples that are replicated; and 941 mussel tissue samples, which includes 430 samples that are replicated. Each sample has been analyzed for 63 unique aromatic and alkane hydrocarbon classes simultaneously. These large numbers of analyte classes and of replicated samples make it possible to identify associations of outlier samples at a high level of confidence.

The procedure described below consists of two parts, which are described sequentially. Part I describes the methods for identifying outlier samples, and Part II describes the methods used to examine the distribution of identified outlier samples among artificial predictor variables. Finally, a brief summary

of results is presented in Part III. The artificial predictor variables considered include the identification number of the catalogue (i.e., batch of samples), and the project responsible for sample collection. Individual samples suspected of systematic bias on the basis of these methods are identified in the **QCERROR** column of the **RECOVERY** table in the NRDA hydrocarbon database. Inclusion of samples identified as biased by these methods in a data set may substantially reduce the power of subsequent statistical tests by inflating the estimate of the sample variance, and by distorting the apparent underlying distribution of the hydrocarbon data.

PART I. IDENTIFICATION OF OUTLIER SAMPLES

A. Structure of the data

Samples were assigned to catalogs under a stratified random system where the stratum was priority level; samples with high priority were analyzed first. Priority levels were assigned by investigators and batches with similar priority levels usually consisting of combinations of control and "oiled" samples. No samples were analyzed in more than one catalog, but different matrices could be included in a catalog. Some samples could be associated into replicate groups according the following criteria for replicate: samples collected on the same date under the same project at the same location within an area of less than 1 m^2.

Each investigator contributing samples to the database was asked to identify samples meeting these criteria for replicate groups. Catalogs could contain all or a portion of the samples from a replicate group. Control sample locations were obtained by polling the investigators. All control locations were selected a priori by the investigators.

Outlier samples fail to conform with one or both of the following expectations. First, hydrocarbon concentrations in replicate groups are expected to be more or less similar. Second, samples collected from a priori control sites were not expected to contain hydrocarbons characteristic of Exxon-Valdez crude oil.

B. Replicate Sample Outliers (Type I Deviants)

A sample was considered an outlier if more than 9 of its 63 hydrocarbon classes were simultaneously very different, when compared with respective concentrations in the remaining samples of the replicate group. These outlier samples are referred to as Type I deviants.

The first step in identifying deviant samples in a replicate group is to identify replicate groups that contain outlier samples. For each hydrocarbon, the logarithm of the squared range of values for the hydrocarbon for each replicate group is plotted against the logarithm of the median value for that group. (Replicate groups that have zero range for the hydrocarbon

considered are not included because the samples in the group obviously do not deviate, failing to meet the criterium for a deviant replicated sample.) The log-log plot accounts for the expected increase in the variance of each hydrocarbon at higher concentration. A linear regression line is calculated for this plot, and the replicate groups associated with the highest 5% of positive deviations from the regression line are identified and given a score of 1, indicating that the replicate group's range for a hydrocarbon concentration was deviant. Only positively deviant replicates on the plot are identified because these have the largest ranges; the negatively deviant replicates are those that agree most closely for the hydrocarbon under consideration. Thus, 63 regressions are calculated and replicate groups had scores ranging from 0 to 63. The score indicated the number of times the replicate group had deviant hydrocarbon ranges. These deviant ranges arise from high hydrocarbon observations in some of the samples included in the group.

Replicate groups with a score greater than 9 were subsequently examined to determine which samples in the group were outliers. In order for a sample to be an outlier it had to have at least 10 hydrocarbon observations that were simultaneously "very different". A hydrocarbon observation was considered "very different" if it met two criteria: 1) its magnitude had to be greater than 10 times the method detection limit (MDL) for that analyte, and 2) the magnitude must have been 3 times the magnitude of the highest remaining observation in the replicate

group. If an observation met these criteria, then the sample was given a score of 1. If a sample accumulated a score greater than 9 (the sample contains more than 9 "very different" observations) then that sample was considered an outlier and flagged.

The reason a sample required 10 or more simultaneous outlier hydrocarbon observations to be identified as a Type I deviant follows. Each sample will contain some number **n** of outlier hydrocarbon measurements. If the distribution of these outliers were random among samples, then each hydrocarbon observation has a 5% probability of being identified as an outlier in each sample. The probability, **P**, that a sample will contain **n** deviant observations simultaneously under these assumptions is:

$$1. \quad P = \binom{k}{n} (0.05)^n (0.95)^{k-n}$$

where **k** = 63 is the number of hydrocarbon classes examined for each sample. According to Eq. 1, the probability that more than 9 hydrocarbon observations are simultaneously deviant within a sample is less than 0.2% (**k** = 63, **n** = 10). This means that the above procedure will mis-identify less than 0.2% of the samples as outliers, if instances of outlier hydrocarbon observations are really randomly distributed among samples.

All samples satisfying the above criteria are identified as Type I deviants and flagged in the database. The process is reiterated without the flagged samples until no more Type I deviant samples are identified. Reiteration is necessary because outlier samples causing the largest ranges are discovered first, and bias the slope of the log-log regression line to obscure less dramatic outliers.

C. Control Site Sample Outliers (Type II Deviants)

A second group of outlier samples, termed Type II deviants, is identified by failure to conform with the expectation that samples collected from a priori control sites are not expected to be contaminated with petroleum hydrocarbons. This expectation is based on a hydrocarbon survey of Prince William Sound conducted in 1977 - 1980, which showed intertidal sediments and mussels to be generally free of petroleum hydrocarbons except in localized areas of vessel traffic (Karinen et al, 1992). Control site samples were collected from control sites picked a priori by the principal investigator (PI) for each project. Petroleum hydrocarbon contamination was be considered present in a control site sample if more than 3 hydrocarbon observations in the following hydrocarbon analyte classes were present at greater than 5 times their respective MDLs: fluorenes, dibenzothiophenes, phenanthrenes, chrysenes, and phytane. Any sample collected from an a priori control site that met this criterium was identified as a Type II deviant and flagged.

PART II.

DISTRIBUTION OF OUTLIER SAMPLES AMONG CATALOGS AND PROJECTS

The distribution of type I deviant samples among catalogs and projects is examined based on an approach that is analogous with eq. 1. Given **j** Type I deviant among a total of **J** samples initially considered, the probability **P** that a project or catalog containing **L** samples of which **m** are deviant is:

$$2. \quad P = \binom{L}{m}\left(\frac{j}{J}\right)^{m}\left(1-\frac{j}{J}\right)^{L-m}$$

assuming the underlying distribution of deviant samples among catalogs or projects is random. These probabilities are first calculated for each project, and the plausibility of the observed probabilities was evaluated using a chi-square test. An estimate of chi-square is calculated as:

$$3. \quad \chi^2 = \sum_{i=1}^{h} \frac{((j/J)L_i - m_i)^2}{(j/J)L_i}$$

where **h** is the number of projects considered. If this estimate is higher than the critical value of chi-square at $\alpha = 0.05$ and **h**-2 degrees of freedom, then all the deviant samples associated in the least probable project are flagged as systematically deviant. A new estimate of chi-square is calculated for the

remaining projects, where both **j** and **J** are reduced by the **m** and **L**, respectively, of the excluded project. The new estimate of chi-square is compared with the critical value, and the process is reiterated until the chi-square estimate is less than the critical value. This process is repeated using catalogs. Catalogs that contain improbably large numbers of Type I deviant samples are listed as Type I catalogs (no projects have yet been identified that contain an improbable number of Type I deviant samples).

A similar process is performed on the Type II deviant samples with some modifications. Type II deviant samples are believed to have come from uncontaminated sites, yet they apparently contain hydrocarbons characteristic of petroleum. Improbable associations of Type II deviant samples with samples sites, sample depths, projects, and catalogs, in that order, are examined using the chi-square procedure described in the preceeding paragraph. Some sites and depths, but no projects, have been found to be associated with an improbably large number of Type II deviant samples. Type II deviant samples associated with these identified sites and depths are therefore excluded, and the distribution of the remaining Type II deviant samples among projects and catalogues are examined. Catalogs containing improbably large numbers of Type II deviant samples were listed as Type II catalogs.

PART III. FINAL DATA EVALUATION

Type I Deviants: Examination of replicate group similarity among the sediment samples, using the methods described above, reveals 122 outlier samples. The chi-square analysis identifies 7 catalogs with disproportionate numbers of deviant samples.

Type II Deviants: Examination of the samples taken from control sites reveal that 6 sites (Simpson Bay, Longb, Maclh, Mcclb, Dayvi, and Ugakb) and the 100 m depth contour contain disproportionate numbers of contaminated samples; these sites and depths may have been contaminated prior to the Exxon-Valdez oil spill. The control site sample analysis also reveal 15 catalogs that appear to contain disproportionate numbers of contaminated samples. Samples from replicate groups or control sites do not appear to be biased by the project that collected them.

Tables 1 and 2 summarize the results of these analyses for the sediment and mussel data, respectively. Note that many catalogs do not contain control site samples, and others contain large numbers of unreplicated samples. Catalogs may therefore be classified according to the number of replicated samples and control site samples they contain. Catalogs in the group most amenable to evaluation using the methods described above include samples from at least 5 replicate groups that are replicated outside the catalog, and at least 5% of the samples in the catalog are from control sites; these catalogs are identified as

"fully evaluatable" in tables 1 and 2. A second group of catalogs meet only one of these criteria, and are identified as "partially evaluatable". A third group of catalogs meet meet none of these criteria, and are identified as "marginally evaluatable". Among the sediment samples, 28 catalogs are fully evaluatable, 29 catalogs are partially evaluatable, and 17 catalogs are marginally evaluatable.

Conclusions

Sediments:

Samples in catalogs that contain improbably large numbers of both Type I and Type II deviant samples are considered biased. These samples are labeled "biased" in the **QCERROR** column of the **SAMPLE** table in the database. Four fully evaluatable catalogs meet these criteria: 6471, 6472, 6476, and 6699. Samples in two partially evaluatable catalogs are also labeled as "biased", those in catalogs 6470 and 6474. In catalog 6470, there are too few replicated samples to evaluate, but 8 of 8 control site samples are Type II deviant. In catalog 6474, there are no control site samples, but 22 of 44 samples are Type I deviants, which is extremely improbable.

Samples in remaining catalogs are labeled as either "suspect" or as "good" in the **QCERROR** column of the **SAMPLE** table in the database. Samples in catalogs that contain improbably large

numbers of either Type I or Type II deviant samples are labeled "suspect", otherwise they are labeled "good".

Of the 2,698 sediment samples processed, 252 were labeled "biased", 466 "suspect" and 1980 "good". Samples labeled "biased" should be used with extreme caution for statistical analyses. Samples labeled "suspect" should be used with some caution because there is reason to believe they are biased, but the results are not definitive. Samples labeled "good" do not appear biased on the basis of the methods and criteria described above, although this may be a result of insufficient replicate and control site samples in the catalog.

Mussels:

No biases have been detected in the mussel sample data. There are 12 catalogs in common between the sediment and mussel data. Of these, only catalog 6116 contained suspect data. Catalog 6116 appeared on the Type II list after analyzing the sediment data, and there are no control samples of mussel tissue in the catalog. Mussel samples associated with this catalog are labeled "suspect". All other mussel samples are labeled "good".

Literature Cited

Karinen, J F, Babcock M M, Brown D B, MacLeod W D, Ramos L S, Short J W. 1992. Hydrocarbons in Intertidal Sediments and Mussels from Prince William Sound, Alaska, 1977-1980: Characterization and Probable Sources. NOAA Tech. Memo., in review.

Table 1. Number of samples in each catalog and number of Type I and II deviants, level of evaluability and final QCERROR CODE for catalogs containing sediment samples

Catno	Number of samples in catalog	Number of unreplicated samples in catalog	Number of Type I deviants in catalog	Number of groups in catalog	Number of groups replicated outside catalog	Number of control samples in catalog	Number of Type II deviants in catalog	How catalog is listed	Are there 5% controls?	Level of evaluatability	QCERROR code
6116	18	8	1	10	10	1	1	II	YES	FULL	SUSPECT
6168	52	37	0	15	15	7	1		YES	FULL	GOOD
6472	43	20	15	22	22	12	12	III	YES	FULL	BIASED
6477	40	24	2	16	16	4	0		YES	FULL	GOOD
6478	44	15	2	28	28	7	0		YES	FULL	GOOD
6548	38	13	2	18	14	4	0		YES	FULL	GOOD
6549	46	18	3	16	10	5	1		YES	FULL	GOOD
6582	41	0	6	18	7	6	0	I	YES	FULL	SUSPECT
6584	43	0	0	22	22	12	0		YES	FULL	GOOD
6588	44	0	0	26	25	3	0		YES	FULL	GOOD
6589	40	0	2	23	23	20	5		YES	FULL	GOOD
6592	45	10	1	17	10	10	0		YES	FULL	GOOD
6594	47	0	1	18	7	12	3		YES	FULL	GOOD
6596	45	6	1	16	7	10	5	II	YES	FULL	SUSPECT
6693	47	20	0	23	23	4	2		YES	FULL	GOOD
6696	42	8	0	24	22	3	0		YES	FULL	GOOD
6697	43	4	2	28	28	3	3	II	YES	FULL	SUSPECT
6698	45	8	3	30	27	10	3		YES	FULL	GOOD
6699	43	3	4	29	28	7	5	III	YES	FULL	BIASED
6700	41	10	0	23	22	5	0		YES	FULL	GOOD
6701	40	5	2	26	23	10	10	II	YES	FULL	SUSPECT
6702	44	1	0	31	29	5	3	II	YES	FULL	SUSPECT
6703	29	2	0	15	11	11	2		YES	FULL	GOOD

Table 1, continued.

Catno	Number of samples in catalog	Number of unreplicated samples in catalog	Number of Type I deviants in catalog	Number of groups in catalog	Number of groups replicated outside catalog	Number of control samples in catalog	Number of Type II deviants in catalog	How catalog is listed	Are there 5% controls?	Level of evaluatability	QC ERROR code
6703	29	2	0	15	11	11	2		YES	FULL	GOOD
NMFS_082	32	1	0	16	9	16	0		YES	FULL	GOOD
NMFS_083	26	3	0	12	8	13	0		YES	FULL	GOOD
NMFS_125	21	0	0	12	10	2	0		YES	FULL	GOOD
NMFS_127	15	0	0	8	8	2	0		YES	FULL	GOOD
NMFS_128	17	0	0	8	5	6	0		YES	FULL	GOOD
6231	43	0	1	15	1	12	0		YES	PARTIAL	GOOD
6241	39	0	0	15	4	8	0		YES	PARTIAL	GOOD
6470	35	31	0	3	2	8	8	II	YES	PARTIAL	BIASED
6545	48	48	0	0	0	5	0		YES	PARTIAL	GOOD
6547	48	48	0	0	0	8	0		YES	PARTIAL	GOOD
6550	34	31	0	3	3	7	4	II	YES	PARTIAL	SUSPECT
6580	46	32	0	6	3	16	2		YES	PARTIAL	GOOD
6585	42	0	0	16	3	18	0		YES	PARTIAL	GOOD
NMFS_129	17	0	0	7	2	6	0		YES	PARTIAL	GOOD
NMFS_183	15	13	0	2	2	1	0		YES	PARTIAL	GOOD
NMFS_184	30	30	0	0	0	2	0		YES	PARTIAL	GOOD
6166	43	23	0	11	11	1	0		NO	PARTIAL	GOOD
6167	85	40	0	35	30	0	0		NO	PARTIAL	GOOD
6471	44	4	19	40	40	2	2	III	NO	PARTIAL	BIASED
6474	44	2	22	40	40	0	0	I	NO	PARTIAL	BIASED
6476	43	10	10	33	33	1	1	III	NO	PARTIAL	BIASED

Table 1, continued.

Catno	Number of samples in catalog	Number of unreplicated samples in catalog	Number of Type I deviants in catalog	Number of groups in catalog	Number of groups replicated outside catalog	Number of control samples in catalog	Number of Type II deviants in catalog	How catalog is listed	Are there 5% controls?	Level of evaluatability	QCERROR code
6578	43	5	0	21	18	0	0		NO	PARTIAL	GOOD
6579	32	5	0	12	8	4	0		NO	PARTIAL	GOOD
6581	47	7	1	22	19	2	0		NO	PARTIAL	GOOD
6583	48	3	0	22	17	0	0		NO	PARTIAL	GOOD
6586	46	1	1	23	16	0	0		NO	PARTIAL	GOOD
6587	46	0	1	24	20	1	0		NO	PARTIAL	GOOD
6590	41	0	1	25	25	0	0		NO	PARTIAL	GOOD
6591	44	0	1	31	31	0	0		NO	PARTIAL	GOOD
6595	42	1	3	19	9	0	0	I	NO	PARTIAL	SUSPECT
NMES_078	13	0	0	6	5	0	0		NO	PARTIAL	GOOD
NMES_137	36	0	0	18	16	1	0		NO	PARTIAL	GOOD
NMES_141	12	0	1	8	7	0	0		NO	PARTIAL	GOOD
NMES_174	44	8	0	15	5	9	0		NO	PARTIAL	GOOD
NMES_175	21	6	5	9	8	0	0	I	NO	PARTIAL	SUSPECT
6104	10	10	0	0	0	0	0		NO	MARGINAL	GOOD
6162	15	6	1	3	0	0	0		NO	MARGINAL	GOOD
6165	9	0	1	3	0	0	0		NO	MARGINAL	GOOD
6230	44	0	1	13	4	0	0		NO	MARGINAL	GOOD
6234	43	0	3	15	2	0	0		NO	MARGINAL	GOOD
6240	43	0	2	13	4	1	1	II	NO	MARGINAL	SUSPECT
6546	46	45	0	1	1	1	1	II	NO	MARGINAL	SUSPECT
6593	48	9	0	16	1	0	0		NO	MARGINAL	GOOD

Table 1, continued

Catno	Number of samples in catalog	Number of unreplicated samples in catalog	Number of Type I deviants in catalog	Number of groups in catalog	Number of groups replicated outside catalog	Number of control samples in catalog	Number of Type II deviants in catalog	How catalog is listed	Are there 5% controls?	Level of evaluatability	QCI ERROR code
6593	48	9	0	16	1	0	0		NO	MARGINAL	GOOD
6658	17	17	0	0	0	0	0		NO	MARGINAL	GOOD
6694	49	42	1	4	4	2	2	II	NO	MARGINAL	SUSPECT
6695	45	45	0	0	0	1	1		NO	MARGINAL	GOOD
6713	37	37	0	0	0	0	0		NO	MARGINAL	GOOD
6714	20	20	0	0	0	0	0		NO	MARGINAL	GOOD
NMFS_075	12	0	0	5	3	0	0		NO	MARGINAL	GOOD
NMFS_081	1	0	0	1	1	0	0		NO	MARGINAL	GOOD
NMFS_126	8	0	0	5	4	0	0		NO	MARGINAL	GOOD
NMFS_191	9	1	0	4	4	0	0		NO	MARGINAL	GOOD

Table 2. Number of samples in each catalog and number of Type I and II deviants, level of evaluability and final QCERROR CODE for catalogs containing mussel tissue samples.

Catno	Number of samples in catalog	Number of unreplicated samples in catalog	Number of Type I deviants in catalog	Number of groups in catalog	No. of groups replicated outside catalog	Number of control samples in catalog	Number of Type II deviants in catalog	How catalog is listed	Are there 5% controls?	Level of evaluatability	QCERROR code
6473	27	11	0	13	11	3	0		YES	FULL	GOOD
6475	30	19	0	11	11	3	0		YES	FULL	GOOD
NMFS_084	40	0	0	19	11	12	0		YES	FULL	GOOD
NMFS_085	8	0	0	7	7	2	0	*	YES	FULL	GOOD
NMFS_125	21	0	0	11	10	2	0		YES	FULL	GOOD
NMFS_127	19	0	0	10	10	2	0		YES	FULL	GOOD
NMFS_159	26	0	0	14	9	13	0		YES	FULL	GOOD
6243	35	35	0	0	0	20	1		YES	PARTIAL	GOOD
6244	40	10	0	7	0	5	0		YES	PARTIAL	GOOD
NMFS_027	6	4	0	1	0	1	0		YES	PARTIAL	GOOD
NMFS_030	2	2	0	0	0	1	0		YES	PARTIAL	GOOD
NMFS_031	8	8	0	0	0	1	0		YES	PARTIAL	GOOD
NMFS_077	4	4	0	0	0	3	0		YES	PARTIAL	GOOD
NMFS_099	24	24	0	0	0	12	0		YES	PARTIAL	GOOD
NMFS_124	7	7	0	0	0	1	0		YES	PARTIAL	GOOD
NMFS_126	24	0	0	10	4	6	0		YES	PARTIAL	GOOD
NMFS_128	14	1	0	6	4	4	0		YES	PARTIAL	GOOD
NMFS_129	12	2	0	5	1	3	0		YES	PARTIAL	GOOD
NMFS_136	35	2	0	11	4	2	0		YES	PARTIAL	GOOD
NMFS_147	29	1	0	10	0	9	6		YES	PARTIAL	GOOD
NMFS_185	1	0	0	1	1	1	0		YES	PARTIAL	GOOD
6659	37	6	1	24	17	0	0		NO	PARTIAL	GOOD
6692	72	36	0	18	17	2	0		NO	PARTIAL	GOOD

Table 2 continued.

Catno	Number of samples in catalog	Number of unreplicated samples in catalog	Number of Type I deviants in catalog	Number of groups in catatalog	Number of groups replicated outside catalog	Number of control samples in catalog	Number of Type II deviants in catalog	How catalog is listed	Are there 5%a controls?	Level of evaluatability	QC ERROR code
6104	13	13	0	0	0	0	0		NO	MARGINAL	GOOD
6116	6	0	0	2	0	0	0	II	NO	MARGINAL	SUSPECT
6122	3	3	0	0	0	0	0		NO	MARGINAL	GOOD
6157	9	0	0	3	0	0	0		NO	MARGINAL	GOOD
6159	6	4	0	1	0	0	0		NO	MARGINAL	GOOD
6160	55	53	0	2	2	0	0		NO	MARGINAL	GOOD
6161	57	55	0	2	2	0	0		NO	MARGINAL	GOOD
6162	4	2	0	1	0	0	0		NO	MARGINAL	GOOD
6165	11	2	0	3	0	0	0		NO	MARGINAL	GOOD
6166	9	3	0	2	0	0	0		NO	MARGINAL	GOOD
6167	3	1	0	1	0	0	0		NO	MARGINAL	GOOD
6542	35	35	0	0	0	0	0		NO	MARGINAL	GOOD
6543	47	47	0	0	0	0	0		NO	MARGINAL	GOOD
6544	46	44	0	1	0	0	0		NO	MARGINAL	GOOD
6642	32	32	0	0	0	0	0		NO	MARGINAL	GOOD
6653	1	1	0	0	0	0	0		NO	MARGINAL	GOOD
6691	32	32	0	0	0	0	0		NO	MARGINAL	GOOD
NMFS 028	1	1	0	0	0	0	0		NO	MARGINAL	GOOD
NMFS 032	4	4	0	0	0	0	0		NO	MARGINAL	GOOD
NMFS 033	5	3	0	1	0	0	0		NO	MARGINAL	GOOD
NMFS 089	2	0	0	1	1	0	0		NO	MARGINAL	GOOD
NMFS 141	12	3	0	5	4	0	0		NO	MARGINAL	GOOD

Table 2 continued.

Catno	Number of samples in catalog	Number of unreplicated samples in catalog	Number of Type I deviants in catalog	Number of groups in catalog	Number of groups replicated outside catalog	Number of control samples in catalog	Number of Type II deviants in catalog	How catalog is listed	Are there 5% controls?	Level of evaluatability	QCTRROR code
NMFS_089	2	0	0	1	1	0	0		NO	MARGINAL	GOOD
NMFS_141	12	3	0	5	4	0	0		NO	MARGINAL	GOOD
NMFS_176	28	1	0	13	4	0	0		NO	MARGINAL	GOOD