Nonparametric Identification of Petrogenic and Pyrogenic Hydrocarbons in Aquatic Ecosystems

MARK G. CARLS[†]

National Oceanics and Atmospheric Administration, National Marine Fisheries Service, Alaska Fisheries Science Center, Auke Bay Laboratory, 11305 Glacier Highway, Juneau, Alaska 99801

Novel nonparametric models developed herein discriminated between oiled and nonoiled or pyrogenic and oiled sources better than traditionally used diagnostic ratios and can outperform previously published oil identification models. These methods were compared using experimental and environmental hydrocarbon data (sediment, mussels, water, and fish) associated with the *Exxon Valdez* oil spill. Several nonparametric models were investigated, one designed to detect petroleum in general, one specific to Alaska North Slope crude oil (ANS), and one designed to detect pyrogenic PAH. These ideas are intended as guidance; nonparametric models can easily be adapted to fit the specific needs of a variety of petrogenic and pyrogenic sources. Oil identification was clearly difficult where composition was modified by physical or biological processes; model results differed most in these cases, suggesting that a multiple model approach to source discrimination may be useful where data interpretation is contentious. However, a combined nonparametric model best described a broad range of hydrocarbon sources, thus providing a useful new analytical assessment tool.

Introduction

About 1.3 million metric tons of petroleum from natural seeps, oil spills, and other human activities enter the world's oceans each year (*1*). Determining natural and human sources of hydrocarbon contamination in the environment is necessary to understand why they are present, who (or what) might be responsible, and the biological and ecological implications of their presence.

The presence of polynuclear aromatic hydrocarbons (PAHs) in water, sediment, and tissue provides clues regarding contaminant sources, although identification can be complicated by composition changes as the oil is altered by physical, chemical, and biological processes. Not only can PAHs be used for source identification, they are likely primarily responsible the toxicity of petroleum, adding to the incentive to quantify them (*2–4*). Numerous diagnostic PAH ratios have been used to discriminate among sources (*5*). Frequently, these ratios are examined to discriminate between petrogenic sources and others, such as pyrogenic sources.

More complex models were developed to distinguish *Exxon Valdez* oil, an Alaska North Slope crude oil (ANS), from other hydrocarbons in Prince William Sound (*6, 7*).

[†] Phone: (907)789-6019; fax: (907)789-6094; e-mail: mark.carls@noaa.gov.

Both models rely on analysis of PAH composition, and both were intended primarily for situations where whole oil was present, such as oil on intertidal sediment. The oil-fingerprinting method (OFM; *6*) relies on a series of decisions for identification, including the presence or absence of specific PAH and values of PAH ratios. The first-order loss-rate model (FORLM; *7*) relies on first-order loss-rate constants for 14 environmentally persistent PAH and bootstrapped error distributions to determine if composition is consistent with ANS. Results provided by these models are not identical, and rather than arbitrarily choosing one or the other, an independent perspective could provide additional useful information.

The purpose of this study is to present an alternative, nonparametric approach to source identification. This method is distinct because it relies on scoring, not specific concentrations. Two nonparametric models were developed to identify petrogenic PAH and one to identify pyrogenic PAH. The nonspecific version (PSCORE$_{oil}$) is simply based on the presence or absence of homologous PAH families and their constituents. Scoring in the ANS-specific version (PSCORE$_{ANS}$) is based on the observation that fewer unsubstituted parent compounds are present in ANS than alkyl-substituted compounds in each of five homologous families (naphthalenes, fluorenes, dibenzothiophenes, phenanthrenes, and chrysenes). This relationship is generally true in unweathered ANS (except for C3-fluorenes, C4-phenanthrenes, and C4-chrysenes) and remains (or becomes) true as the oil weathers. The same relationship is evident in Kuroshima Bunker oil, Montreal Bunker oil, and Katalla crude oil (*8*). The nonparametric pyrogenic model is based on the observation that unsubstituted parent homologue concentrations are typically much greater than alkyl-substituted concentrations.

Experimental data were first examined with all models and ratios to verify and compare performance. Scores that distinguished oil in the nonparametric models were determined empirically. The OFM and FORLM model results were each subdivided into nonspecific and ANS-specific outputs to compare to analogous nonparametric results. The source-discrimination ability of all models and ratios were then evaluated in environmental data, resulting in several combined models. A consensus model combined results of PSCORE, OFM, and FORLM; this was extended to include two pyrogenic estimates. A third combined model was based solely on nonparametric petrogenic and pyrogenic models.

Materials and Methods

Postspill PAH data from Prince William Sound, Alaska, were obtained from the Natural Resource Damage Assessment database (*9*). Included were 1907 mussel samples, 1505 sediment samples, 236 water samples, and 290 fish samples, combined across 19 studies. Low-density polyethylene passive samplers served as surrogate tissue in 3 additional studies. Experimental data included ANS, water, sediment, and fish tissue, assembled across seven studies.

Concentrations of PAH in water, sediment, tissue, and low-density polyethylene passive samplers were determined by gas chromatography/mass spectroscopy after addition of perdeuterated hydrocarbon standards and extraction with dichloromethane (*10*; Supporting Information 1). Hydrocarbons were isolated and purified by silica gel/alumina column chromatography followed by size-exclusion high-performance liquid chromatography and fractionation; water samples were only processed through aluminum/silica columns. Method detection limits depended on sample

weights and generally were 1 ng/g in sediment or tissue and $1-8$ ng/L in water. Accuracy of the hydrocarbon analyses was about $\pm 15\%$ based on comparison with National Institute of Standards and Technology values; precision, expressed as the coefficient of variation, was usually less than about 20%. Concentrations less than method detection limits were treated as zero except as noted. Total PAH (TPAH) is the sum of 39 PAH; naphthalenes (N0–N4), biphenyl (BPH), acenaphthylene (ACY), acenaphthene (ACE), fluorenes (F0–F3), dibenzothiophenes (D0–D3), phenanthrenes (P0–P4), anthracene (ANT), fluoranthene (FLA), pyrene (PYR), C1-fluoranthenes/pyrenes (C1FLA), benzo[*a*]anthracene (BaA), chrysenes (C0–C4), benzo[*b*]fluoranthene (BbF), benzo[*k*]-fluoranthene (BkF), benzo[*e*]pyrene (BeP), benzo[*a*]pyrene (BaP), perylene (PER), indeno(123-*cd*)pyrene (IDP), dibenzo-(*a,h*)anthracene (DBA), and benzo(*ghi*)perylene (BZP).

The nonspecific PSCORE$_{oil}$ model was developed to detect petroleum hydrocarbons. Possible scores ranged from 0 to 5, determined as follows. For a maximum of 2.5 points, the score was increased by 0.5 for each of five homologous families present in ANS (N, F, D, P, and C) where the TPAH concentration was >0. An additional 0.5 was added for each family if more than one N, F, or D homologue was present or if more than two P or C homologues were present (for a maximum of 2.5 additional points). This latter scoring was based on the presence of homologues in weathered *Exxon Valdez* oil: three N, F, and D homologues were often present, but the proportion of the least-substituted homologues was very low. At least three P homologues were well represented in the ANS profile and four C homologues were present. However, detection of chrysenes was often relatively poor because typical concentrations were relatively small; hence, just three C homologues were required to increment the score. Determination of successful oil identification by PSCORE and accurate rejection of samples without oil was accomplished empirically with experimental data to determine the score that accomplished both objectives.

The ANS-specific version of PSCORE relied on the observation that $X_{0j} < X_{ij}$ in ANS, where X_{ij} represents each homologous family (N, F, D, P, C); $j = 1-5$, $i = 1-n_j$, and n_j = no. of homologues examined in the *j*th family (4, 3, 3, 4, and 1, respectively). The score was incremented each time the relationship was true, e.g., when $X_{0j} < X_{ij}$ the score was incremented by $1/n_j$. Division by n_j allowed equal weighting of each homologous family. Chrysenes were restricted to comparison of C0 and C1 because more-substituted chrysenes were not always detected. Scores ranged from $0-5$ and were compared directly to PAH diagnostic ratio results. To relate model results to the OFM and FORLM, an empirically determined value was used to discriminate the presence or absence of oil as explained above.

The nonparametric pyrogenic detection model was based on the observation that $X_{0j} \gg X_{ij}$ in PAH from pyrogenic sources. Weighed scores within homologues were assigned as follows: if $X_{0j} > 10X_{ij}$ then $s_{ij} = 1/n_j$, if $X_{0j} > 5X_{ij}$ then $s_{ij} = 0.5/n_j$, if $X_{0j} > 2.5X_{ij}$ then $s_{ij} = 0.25/n_j$, if $X_{0j} > X_{ij}$ then $s_{ij} = 0.1/n_j$ where s_{ij} = the *i*th score in the *j*th homologous family (N, F, D, P, C). The pyrogenic score was the sum of s_{ij}. As before, chrysenes were restricted to comparison of C1 and C0.

The FORLM required the presence of 14 environmentally persistent PAH (N3, N4, F2, F3, D1–D3, P1–P4, and C0–C2) to yield a probability estimate that ANS was (or was not) present. Bootstrapped error distributions from experimental and environmental samples provided the basis for testing the null hypothesis that PAH composition was consistent with that of weathered ANS (*7*). Two model outputs are presented in this paper, source-specific ANS detection and the number (or percentage) of estimable samples (the presence of all 14 required PAH is interpreted in this paper

as nonspecific detection of petroleum).

Detection of ANS by the OFM required the presence of alkylated-N, D2, and C2. Results were classified as provisional diesel or water-soluble fractions of oil if C2 was absent. Although the model uses C2/P2 and D2/P2 ratios to further distinguish among possible sources, the authors indicated that all crudes and crude–diesel mixtures identified before these latter tests be classified as ANS in these data (*6*, *9*). Nonspecific detection of petroleum hydrocarbons was interpreted as petrogenic detection (OFM$_{oil}$). Specific identification of ANS is described as OFM$_{ANS}$.

Model Performance in Controlled Laboratory Samples. PAH compositions in ANS and in water, sediment, and fish tissue exposed to ANS (plus controls) were drawn from experimental studies completed between 1991 and 2001 (Supporting Information 1). Water (saltwater or saltwater–freshwater mixtures) was contaminated by passage through oiled-rock columns (e.g., *2*, *19*) or by direct oil–water mixing (*20*). Fish tissues were Pacific herring (*Clupea pallasi*) muscle (8–18 days of exposure to oiled-rock column effluent) and pink salmon (*Oncorhynchus gorbuscha*) whole fry and fry carcasses (sans head and viscera, 42 days of exposure to oiled food; Supporting Information 1). Samples were classified either as oiled or control. Model performance in control samples provided a measure of false positives. Performance in oiled groups indicated how successful each model was in identifying the presence of known oil contamination.

Model Performance in Environmental Samples. The compositions of PAH in water, sediment, fish, and mussels (*Mytilus trossulus*) from Prince William Sound after the *Exxon Valdez* oil spill were examined. The working hypothesis was that petroleum was increasingly likely in environmental samples as TPAH concentration increased. Tissues representative of whole fish were analyzed, mostly pink salmon and Pacific herring (87%; see Supporting Information 2 for other species).

The ability of each model to identify oil in environmental samples was evaluated by logistic regression. Percentages of samples where oil or ANS were detected were calculated over discrete, logarithmically increasing TPAH concentration intervals. In each interval and for each model, percentages of samples with oil were calculated by dividing the number of oiled samples detected by the total number of samples in that interval.

Diagnostic PAH Ratios. Diagnostic ratios were computed for environmental mussel and sediment samples and included several that increase when the source is petrogenic and several that increase when the source is pyrogenic. Petrogenic models include C0/BaA (*11*), P1/P0 (e.g., *12*), P0/ANT (*13*), the fossil fuel index $[(\sum(N0...N4) + \sum(D0...D3) + 0.5(P0 + P1) + \sum(P2...P4))/TPAH]$ (*14*), and low/high-molecular weight PAH $[(P0 + ANT + PYR + FLA)/(BaA + C0 + BbF + BkF + BaP + BeP + PER + IDP + DBA + BZP]$ (*11*). Pyrogenic models include the pyrogenic fraction $[ACE + ACN + ANT + FLA + PYR + BaA + C0 + BbF + BkF + BeP + BaP + IDP + DBA + BZP)/TPAH]$ (*15*), FLA/PYR (*13*), percent perylene, pyrogenic index $[(ACN + ACE + ANT + FLU + PYR + BaA + BbF + BkF + BeP + BaP + PER + IDP + DBA + DBZ)/(\sum(N0...N4) + \sum(F0...F3) + \sum(D0...D3) + \sum(P0...P4) + \sum(C0...C3)]$ (*5*), (FLA + PYR)/(P2 + P3) (*16*), (FLA + PYR)/$\sum$(P1...P4) (*17*), and $\sum$(C1...C4)/C0 (*18*).

To provide a model-free summary of trends as a function of TPAH concentration, mean ratios were computed within increasing TPAH concentration groups (0.25 log intervals), where TPAH ranged from 0 to about 10^2 μg/L or 10^7 ng/g dry weight, depending on sample type. Means were smoothed with a combination of running medians and weighted averaging to describe sample trends (4253H; *21*).

Initial analysis demonstrated loss of detail for pyrogenic ratios when data were corrected by method detection limits;

TABLE 1. Detection of Petroleum Hydrocarbons in Laboratory Controls (A) and Samples Experimentally Oiled (B and C) with Alaska North Slope Crude Oil (ANS) by Three Models, Nonparametric Model PSCORE, a PAH Oil-Fingerprint Model (OFM; 6), and a PAH First-Order Loss-Rate Model (FORLM; 7)[a]

(A) Laboratory Controls

matrix	TPAH			PSCORE		OFM		FORLM	
	mean	SE	n_{total}	% oil	% ANS	% oil	% ANS	% oil	% ANS
water	0.04	0.01	27	0.0	0.0	0.0	0.0	0.0	0.0
sediment	27.6	13.1	24	8.3	0.0	29.2	0.0	0.0	0.0
fish tissue	89.0	6.8	28	0.0	0.0	0.0	0.0	0.0	0.0

(B) Experimentally Oiled with ANS

matrix	TPAH			PSCORE		OFM		FORLM	
	TPAH	SE	n_{total}	% oil	% ANS	% oil	% ANS	% oil	% ANS
pure oil	8283364	2414037	11	100.0	100.0	100.0	100.0	100.0	100.0
sediment	22806	5527	95	93.7	93.7	100.0	97.9	75.8	75.8
water	12	3	104	88.5	57.7	94.2	27.9	27.9	23.1
fish tissue	13918	2325	183	48.6	6.0	44.8	2.2	1.1	1.1

(C) Experimentally Oiled with ANS, Where TPAH Concentration Was > the Control 95% Confidence Bound

matrix	TPAH			PSCORE		OFM		FORLM	
	TPAH	SE	n_{total}	% oil	% ANS	% oil	% ANS	% oil	% ANS
sediment	24904	599	87	100.0	100.0	100.0	97.7	81.6	81.6
water	13	3	94	96.8	62.8	100.0	30.9	30.9	25.5
fish tissue	15244	2524	167	53.3	6.6	49.1	2.4	1.2	1.2

[a] Matrixes examined included pure oil, water, sediment, and fish tissue (Pacific herring and pink salmon). Mean total polynuclear aromatic hydrocarbon (TPAH) concentrations and standard error (SE) are listed for each matrix (μg/L in water, ng/g dry weight in other matrixes). The total number of samples is n_{total}. Section C repeats the analysis in section B except that oiled samples with TPAH concentrations indistinguishable from those of the controls were excluded.

thus, results of all pyrogenic ratios and models are based on raw data.

Combined Models. A consensus petrogenic model (S_e) was assembled from all petrogenic models (PSCORE, OFM, FORLM) and included nonspecific model results (score 0−3) and ANS-specific results (score 4−6). For example, $S_e = 3$ when all nonspecific results were significant and no specific results were significant; $S_e = 5$ for any two significant ANS-specific model results, regardless of nonspecific results. Consensus model S_e was used extensively to interpret diagnostic ratio results; a parallel version that did not include nonparametric results was used to interpret nonparametric model scores (S_{ep}).

An extended consensus model (S_T) that included assessment of two pyrogenic and all petrogenic model outcomes provided a summary of dominant source characteristics: $S_T = S_e - (S_y + S_z)$. Two pyrogenic scores, S_y and S_z, corresponding to (FLA + PYR)/$\sum$(P1...P4) and the nonparametric pyrogenic model, ranged from 0 to 3. Pyrogenic scores = 1 for values that included no more than 10% of samples identified as petrogenic ($S_e > 2$), 2 for 5%, and 3 for 1% petrogenic inclusion. Summary score S_T ranged from −6 to 6: values <0 were progressively more likely to be pyrogenic; those >0 were progressively more likely to be oiled.

A combined model that included only nonparametric results was similarly constructed: $S_N = P_{petro} - P_{pyro}$, where P_{petro} and P_{pyro} were scores based on empirically determined 0.1−20% threshold levels. For example, $P_{petro} = 6$ where $PSCORE_{ANS} = 3.05$; 0.1% of nonoiled scores exceeded this threshold level (Supporting Information 3).

Results and Discussion

Model and Ratio Performance in Controlled Laboratory Samples. Oil was rarely suggested by any of the models in control samples (Table 1, section A). Nonparametric scores that distinguished between oiled and control samples were >3.5 for $PSCORE_{oil}$ and >3.0 for $PSCORE_{ANS}$. All apparent false positive controls were due to 3 of 24 consecutive

sediment samples from a single experiment (TPAH > 50 ng/g in these). The incidence of false positives in ANS-specific models was 0%.

All models readily detected oil in experimental samples where the composition was about the same as in ANS; detection was more difficult in samples where PAH composition was modified by physical or biological processes (Table 1, section B). All models detected 100% oil in pure ANS samples. Oil was readily detected in sediment (76−98% by ANS-specific models). Oil was detected less frequently in water (25−63%) and fish tissue (1−7%) by ANS-specific models. The FORLM was consistently the most conservative model. More oil was detected by generalized models than ANS-specific models in samples where PAH composition was modified physically or biologically. Oil detection by all models was slightly greater when analysis of experimentally oiled samples was limited to those with TPAH concentrations greater than the upper 95% confidence bound of corresponding controls (Table 1, section C).

Petrogenic ratios were usually greater in experimentally oiled samples than in control samples; the reverse was true for pyrogenic ratios (Supporting Information 4). Petrogenic ratios in experimentally oiled samples were typically significantly higher than in control samples ($0.001 < P \le 0.013$). Exceptions were that fossil fuel index values did not differ significantly between oiled and control water ($P = 0.593$) and the low- to high-molecular weight ratio was greater in control sediment ($P = 0.001$). Pyrogenic ratios in experimentally oiled samples were significantly lower than in control water and sediment samples ($P < 0.001$). Pyrogenic ratios in oiled tissue were typically less than in controls (five of six cases) but differences were significant only for the pyrogenic fraction and the pyrogenic index.

Model and Ratio Performance in Environmental Samples. Oil detection in water, sediment, fish, and mussels by each model generally increased sigmoidally with TPAH concentration (Figure 1). In general, detection of oil by the models increased from 0% at low TPAH concentrations to
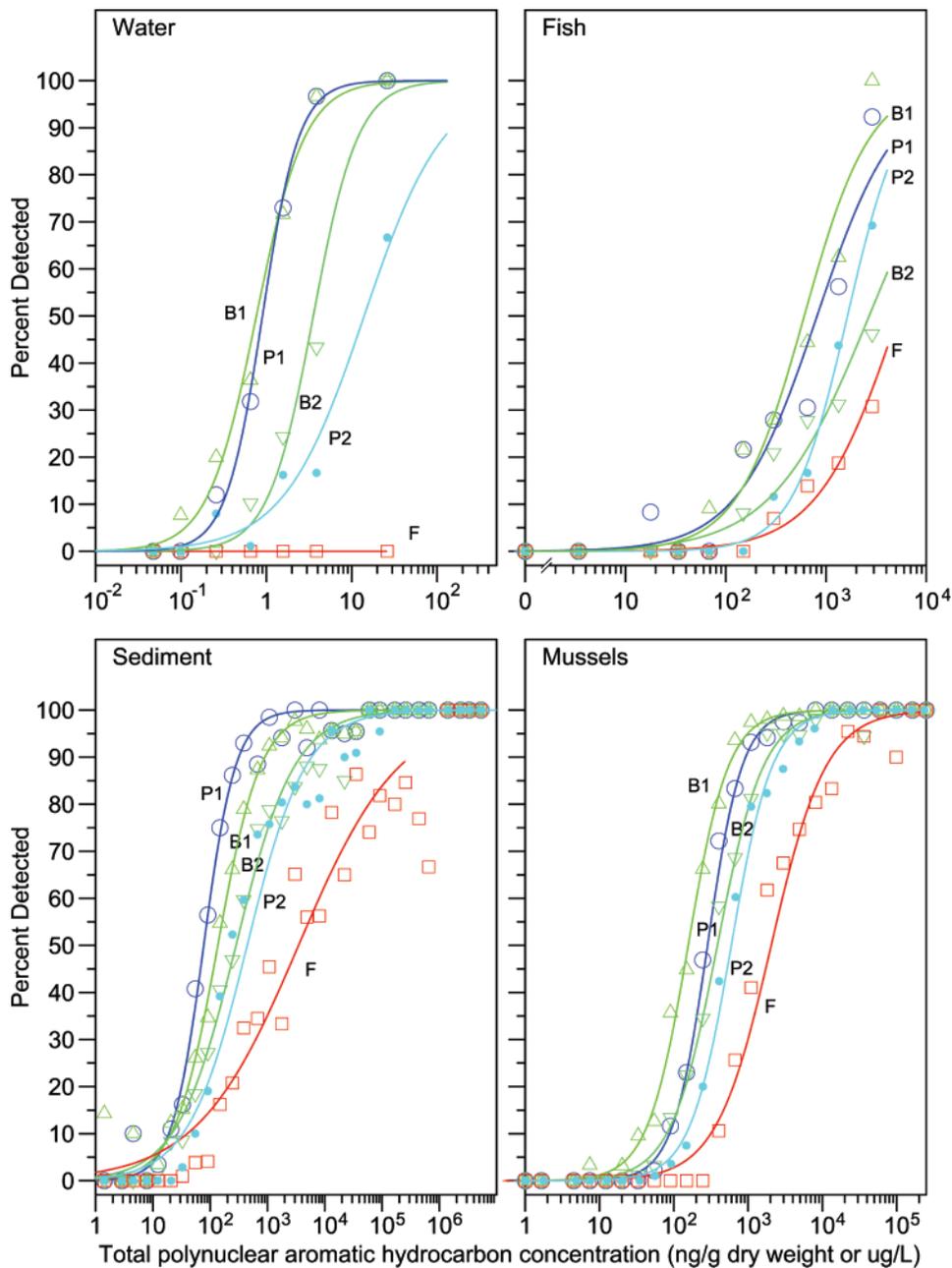
FIGURE 1. Detection of oil in water, fish tissue, sediment, and mussels from Prince William Sound by nonspecific models PSCORE$_{oil}$ (P1) and the OFM$_{oil}$ (B1) and ANS-specific models PSCORE$_{ANS}$ (P2), OFM$_{ANS}$ (B2), and the FORLM$_{ANS}$ (F). Not illustrated are the nonspecific FORLM$_{oil}$ results, which were highly similar to the FORLM$_{ANS}$ results. Plotted curves are logistic regressions.

50−100% at high TPAH concentrations in every matrix (water, sediment, fish, and mussels; Figure 1). Correlation between TPAH and percent oil detection was strong in water ($0.86 \leq r \leq 0.96$), sediment ($0.87 \leq r \leq 0.96$), and mussels ($0.85 \leq r \leq 0.87$) and moderate in fish ($0.56 \leq r \leq 0.66$) with the exception that ANS was typically not detected by the FORLM in water from Prince William Sound because chrysenes were often below detection limits. Detection by nonspecific models (PSCORE$_{oil}$ and OFM$_{oil}$) was consistently greater than by ANS-specific models, and the FORLM was consistently the most conservative model.

Few petrogenic diagnostic ratios adequately discriminated between ANS and other sources (3 of 12; Figure 2), where "adequate" is defined as detection of $\geq 80\%$ of the oiled samples ($S_e > 2$) and that no more than 10% of nonoiled samples ($S_e = 0$) were above a threshold value. Petrogenic ratios generally increased with TPAH concentration in mussel and sediment samples, but relatively few estimates were possible for some ratios (e.g., C0/BaA, P0/ANT) because divisor concentrations were frequently zero (Figure 2 and Supporting Information 5). Scatter tended to be high for some ratios (e.g., the low- to high-molecular weight ratio and the fossil fuel index). Paradoxically, upward trends with increasing TPAH reversed in some cases (P1/P0, P0/ANT, and the low- to high-molecular weight ratio in sediment). Variable success of diagnostic ratios has been noted previously, and results vary among different environmental situations (22).

In contrast to the diagnostic ratios, nonparametric score discrimination between nonoiled and oiled samples was good (PSCORE$_{ANS}$); 93% of oiled mussel and sediment scores ($S_{ep} > 2$) were above a threshold score (2.2−2.4) exceeded by just 1% of the samples least likely to contain oil ($S_{ep} = 0$; Figure 2; Supporting Information 5−6). PSCORE$_{ANS}$ scores increased in sigmoid fashion as TPAH concentration increased, con-
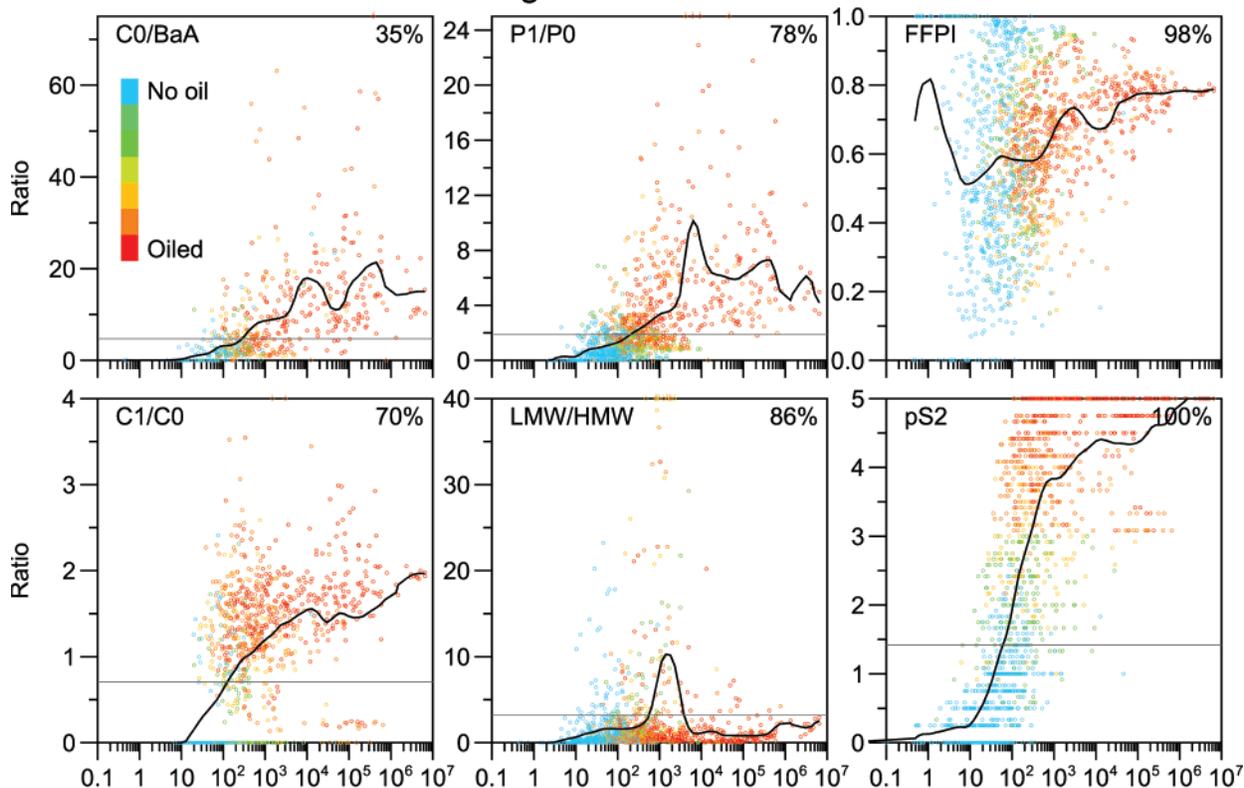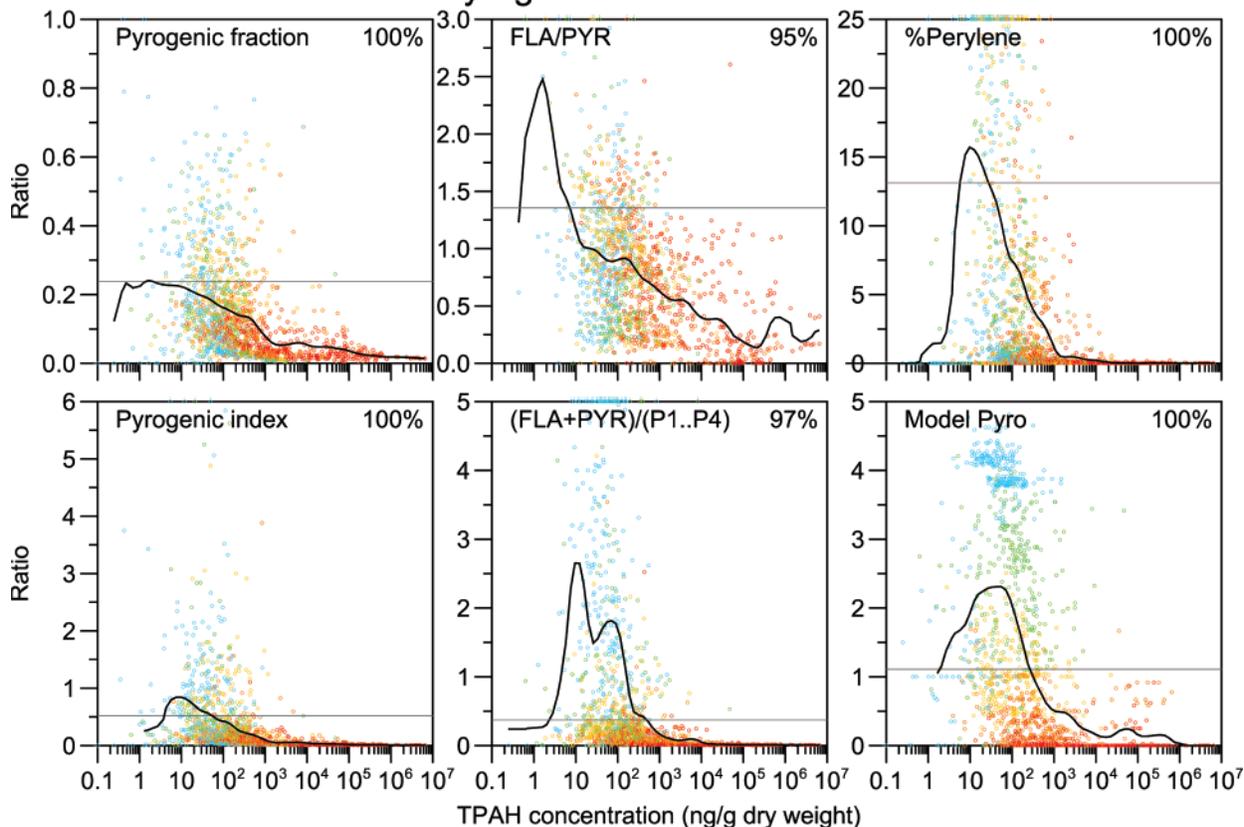
FIGURE 2. Relationship between TPAH and ratios and scores in sediment samples collected in Prince William Sound after the *Exxon Valdez* oil spill. The curves are smoothed diagnostic mean ratios, grouped by increasing TPAH concentration (0.25 intervals in log space). Oil classification (see color key) is based on combined model results, S_e, or S_{ep} (for nonparametric graphs). The horizontal lines are 10% threshold levels (i.e., 90% of the nonoiled samples (blue) are below this line in the top six graphs and 90% of the oiled samples (orange to red) are below the line in the bottom six graphs). FFPI is the fossil fuel index; LMW/HMW is the low- to high-molecular weight ratio; pS2 is PSCORE_ANS. Percentages indicate estimable samples.

TABLE 2. Comparison of Model Results in Samples Containing Various Oils, Pyrogenic Hydrocarbons, and Other Sources, Including Alaska North Slope Crude Oil (ANS), the Water-Accommodated Fraction (WAF) of ANS, and ANS Chemically Dispersed with Corexit 9527[a]

matrix n	ANS[b,c] sediment 5	ANS WAF[d] water 7	ANS dispersed[d] water 8	diesel fuel[e] PEMD 12	Katalla seep oil[f] PEMD 8	Constantine background[g] sediment 16	Monterey oil[g] pure 11	Selendang bunker oil[f] PEMD 26	coal[h] pure 1	creosote[c,h,i] water,pure 3	coal tar[h] pure 1	soot and dust[c] particulate 3
TPAH (ppb)	ng/g	µg/L	µg/L	ng/g	ng/g	ng/g	ng/g	ng/g	ng/g		ng/g	ng/g
min	3705	8.5	2.3	154	581	494	11800	260	526	447 µg/L	103000	249
max	401000	286	469	35500	17700	726	1420000	463000		303000 ng/g		70000
percent scored as petrogenic (nonspecific models)												
FORLM$_{oil}$	100	100	88	42	63	94	100	58		0		33
OFM$_{oil}$	100	100	100	100	75	100	100	100		100		100
PSCORE$_{oil}$	100	100	100	100	100	100	100	100		100		100
percent scored as ANS												
FORLM$_{ANS}$	100	100	75	17	0	6	45	50		0		0
OFM$_{ANS}$	100	100	88	58	63	100	100	58		67		100
PSCORE$_{ANS}$	100	100	38	42	100	88	100	96		0		0
percent scored as pyrogenic												
pyrogenic model	0	0	0	0	0	0	0	0		100		0
(fla + pyr)/Σ(p1...p4)	0	0	0	0	0	0	0	0		67		67
summary scores												
consensus petrogenic (S_e)	6.0	6.0	4.9 (2−6)	3.8 (2−6)	4.6 (4−5)	4.9 (4−6)	5.5 (5−6)	5.1 (4−6)	6.0	3.3 (2−4)	4.0	4.0
consensus pyrogenic and petrogenic (S_T)	6.0	6.0	4.9 (2−6)	3.8 (2−6)	4.6 (4−5)	4.9 (4−6)	5.5 (5−6)	5.1 (4−6)	6.0	−2.3 (−3 to −2)	−2.0	1.3 (1−2)
combined nonparametric model (S_N)	6.0	6.0	5.0 (2−6)	4.4 (0−6)	6.0	5.8 (4−6)	6.0	6.0 (5−6)	6.0	−5.0	−5.0	4.0 (3−5)

[a] Matrixes sampled include pure product, sediment, water, and low-density polyethylene passive sampler devices (PEMDs). The greater the score magnitude, the more likely the identified characteristic, e.g., for model S_T, values <0 were progressively more likely to be pyrogenic; those >0 were progressively more likely to be petrogenic. [b] Fresh to heavily weathered. [c] Ref 23. [d] Ref 20. [e] Ref 25. [f] Ref 27. [g] Ref 9. [h] Ref 24. [i] Ref 26.

sistent with the increasing probability that ANS was present in mussel and sediment samples.

Pyrogenic hydrocarbons were present in Prince William Sound mussels and sediment at low concentrations (Figure 2). Pyrogenic ratios were generally at maximum around 100 ng/g dry weight in mussels and near 10 ng/g in sediment and faded away at lower concentrations (Figure 2, Supporting Information 5). Scatter was often high. Only one of the pyrogenic diagnostic ratios adequately discriminated between *Exxon Valdez* oil and the pyrogenic background in both mussels and sediment. Where no more than 10% of oiled samples were above the threshold value, 79−80% of nonoiled mussel and sediment samples were identified as pyrogenic by the ratio (FLA + PYR)/ Σ(P1...P4). No other ratios passed this test.

The nonparametric pyrogenic model provided the best overall discrimination between nonoiled and oiled samples; 67% and 81% of nonoiled mussel and sediment samples, respectively, were above the threshold score exceeded by just 1% of the samples most likely to contain oil (S_{ep} > 2; Supporting Information 7). Although the values used to parametrize the pyrogenic model were arbitrary, discrimination was highly similar in trials with other coefficients, all requiring that $X_{0j} \gg X_{ij}$ (data not illustrated). However, discrimination declined as the requirement that X_{0j} be much greater than X_{ij} was relaxed, and discrimination was poorer in a trial that simply required that $X_{0j} > X_{ij}$ (80% detection with 10% false positives in sediment). Discrimination was inadequate for a nonparametric pyrogenic model that required $X_{0j} > X_{1j} > X_{2j} > ... X_{nj}$ (22% detection with 10% false positives; data not illustrated).

Detection failure at low concentrations and masking of pyrogenic signals by oil at higher concentrations are the most likely reasons that pyrogenic hydrocarbon concentrations clustered between about 10 and 100 ng/g. *Exxon Valdez* oil

apparently masked the pyrogenic signature at TPAH concentrations >100 ng/g, an inference consistent with Page et al. (17). Detection, hence ability to determine sources, consistently faded away for ratios and models as TPAH concentrations approached zero, even when raw data were examined.

Few samples had both high pyrogenic and high petrogenic scores. For example, TPAH < 200 ng/g in 82% of sediment where pyrogenic scores >1.5 and TPAH > 200 ng/g in 83% of sediment where the petrogenic score > 3. Similarly, TPAH < 200 ng/g in 72% of mussels where pyrogenic scores >1.5 and TPAH > 200 ng/g in 97% of mussels where the petrogenic score > 3.

All models provided strong evidence of oil in Prince William Sound sediment and mussels and weaker evidence of oil in matrixes where physical or biological processes caused changes in PAH composition. The FORLM (7) was consistently the most conservative of all models. Division of model outputs into nonspecific and ANS-specific components provided an advantage; nonspecific models were more likely to adequately detect petroleum in matrixes where the PAH composition was significantly modified by physical or biological processes (e.g., water and fish tissue). The nonspecific models also had distinctly lower detection thresholds than the ANS-specific models. However, nonspecific model results must be interpreted carefully; as will be demonstrated shortly, lack of specificity hampers discrimination.

Combined Models and Applicability to Other Situations. Detection of oil was generally similar but not identical among models, suggesting one reasonable approach is to consider all models simultaneously when interpreting data. Such consensus models may be particularly useful in situations where interpretation is contentious (such as after the *Exxon Valdez* oil spill). The mean consensus petrogenic score (model S_e) increased sigmoidally with TPAH concentration (data not

illustrated) just as for each component model, consistent with the hypothesis that the presence of oil in Prince William Sound samples was increasingly likely as concentration increased.

The combined nonparametric model S_N best described a broad range of hydrocarbon sources, including several petrogenic sources, one possibly mixed source (soot), and two pyrogenic sources (9, 20, 23–27; Table 2). Discrimination among petrogenic and pyrogenic sources by model S_T was hampered by failure of its component models. In particular, nonspecific single models generally discriminated poorly among petrogenic and pyrogenic sources. Furthermore, the OFM$_{ANS}$ generally failed to distinguish pyrogenic from petrogenic sources; hence, S_T performed poorly with some pyrogenic samples (e.g., creosote). Both combined models, however, performed similarly in Prince William Sound samples; 92–93% of mussel and sediment samples identified as oiled by S_N were also identified as oiled by S_T. The nonparametric modeling approach clearly provides a new and useful assessment tool, one with broad applicability.

Acknowledgments

Supporting Information Available

Glossary of terms, additional method and results details, and a figure presenting ratio and model results for mussels. This material is available free of charge via the Internet at http://pubs.acs.org.

Literature Cited

(1) NRC. *Oil in the Sea III: Inputs, Fates, and Effects*; National Research Council (U.S.); Committee on Oil in the Sea: Inputs Fates, and Effects; National Academy of Sciences: Washington, DC, 2003; pp 1–280.

(2) Heintz, R.; Short, J. W.; Rice, S. D. Sensitivity of fish embryos to weathered crude oil: Part II. Increased mortality of pink salmon (*Oncorhynchus gorbuscha*) embryos incubating downstream from weathered *Exxon Valdez* crude oil. *Environ. Toxicol. Chem.* **1999**, *18*, 494–503.

(3) Sundberg, H.; Ishaq, R.; Akerman, G.; Tjarnlund, U.; Zebuhr, Y.; Linderoth, M.; Broman, D.; Balk, L. A bio-effect directed fractionation study for toxicological and chemical characterization of organic compounds in bottom sediment. *Toxicol. Sci.* **2005**, *84*, 63–72.

(4) Incardona, J. P.; Carls, M. G.; Teraoka, H.; Sloan, C. A.; Collier, T. K.; Scholz, N. L. Aryl hydrocarbon receptor-independent toxicity of weathered crude oil during fish development. *Environ. Health Perspect.* **2005**, *113*, 1755–1762.

(5) Wang, Z.; Fingas, M. F. Development of oil hydrocarbon fingerprinting and identification techniques. *Mar. Pollut. Bull.* **2003**, *47*, 423–452.

(6) Bence, A. E.; Burns, W. A. ASTM STP 1219; Fingerprinting hydrocarbons in the biological resources of the *Exxon Valdez* spill area. 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; pp 84–140.

(7) Short, J.; Heintz, R. Identification of *Exxon Valdez* oil in sediments and tissues of PWS. *Environ. Sci. Technol.* **1997**, *31*, 2375–2384.

(8) Short, J. W. Oil identification based on a goodness-of-fit metric applied to hydrocarbon analysis results. *Environ. Forensics* **2002**, *3*, 349–355.

(9) Short, J. W.; Heintz, R. A.; Nelson, B. D.; Maselko, J. M.; Kendziorek, M. F.; Carls, M. G.; Korn, S. *Exxon Valdez Oil Spill of 1989: State/Federal Trustee Council Hydrocarbon Database 1989–1995*; National Oceanic and Atmospheric Administration, National Marine Fisheries Service: Auke Bay, AK, 1996.

(10) Short, J. W.; Jackson, T. J.; Larsen, M. L.; Wade, T. L. Analytical methods used for the analyses of hydrocarbons in crude oil, tissues, sediments, and seawater collected for the natural resource damage assessment of the *Exxon Valdez* oil spill. *Am. Fish. Soc. Symp.* **1996**, *18*, 140–148.

(11) Socolo, H. H.; Garrigues, P. H.; Ewald, M. Origin of polycyclic aromatic hydrocarbons (PAHs) in coastal marine sediments: case studies in Cotonou (Benin) and Aquitaine (France) areas. *Mar. Pollut. Bull.* **2000**, *40*, 387–396.

(12) Zakaria, M. P.; Takada, H.; Tsutsumi, S.; Ohno, K.; Yamada, J.; Kouno, E.; Kumata, H. Distribution of polycyclic aromatic hydrocarbons (PAHs) in rivers and estuaries in Malaysia: a widespread input of petrogenic PAHs. *Envion. Sci. Technol.* **2002**, *36*, 1907–1918.

(13) Doong, R.; Lin, Y. Characterization and distribution of polycyclic aromatic hydrocarbon contaminations in surface sediment and water from Gao-ping River, Taiwan. *Water Res.* **2004**, *38*, 1733–1744.

(14) Boehm, P. D.; Farrington, J. W. Aspects of the polycyclic aromatic hydrocarbon geochemistry of recent sediments in the Georges Bank region. *Environ. Sci. Technol.* **1984**, *18*, 840–845.

(15) Boehm, P. D.; Page, D. S.; Brown, J. S.; Neff, J. M.; Burns, W. A. Polycyclic aromatic hydrocarbon levels in mussels from Prince William Sound, Alaska, USA, document the return to baseline conditions. *Environ. Toxicol. Chem.* **2004**, *23*, 2916–2929.

(16) Savinov, V. M.; Savinov, T. N.; Matishov, G. G.; Dahle, S.; Næs, K. Polycyclic aromatic hydrocarbons (PAHs) and organochlorines (OCs) in bottom sediments of the Guba Pechenga, Barents Sea, Russia. *Sci. Total Environ.* **2003**, *306*, 39–56.

(17) Page, D. S.; Boehm, P. D.; Brown, J. S.; Neff, J. M.; Burns, W. A.; Bence, A. E. Mussels document loss of bioavailable polycyclic aromatic hydrocarbons and the return to baseline conditions for oiled shorelines in Prince William Sound, Alaska. *Mar. Environ. Res.* **2005**, *60*, 422–436.

(18) Neff, J. M.; Stout, S. A.; Gunster, D. G. Ecological risk assessment of polycyclic aromatic hydrocarbons in sediments: identifying sources and ecological hazard. *Integr. Environ. Assess. Manage.* **2005**, *1*, 22–33.

(19) Carls, M. G.; Rice, S. D.; Hose, J. E. Sensitivity of fish embryos to weathered crude oil: Part 1. Low level exposure during incubation causes malformations and genetic damage in larval Pacific herring (*Clupea pallasi*). *Environ. Toxicol. Chem.* **1999**, *18*, 481–493.

(20) Barron, M. G.; Carls, M. G.; Short, J. W.; Rice, S. D. Photoenhanced toxicity of aqueous phase and chemically dispersed weathered Alaska north slope crude oil to Pacific herring eggs and larvae. *Environ. Toxicol. Chem.* **2003**, *22*, 650–660.

(21) Velleman, P. F.; Hoaglin, D. C. *Applications, Basics and Computing of Exploratory Data Analysis*; Duxbury Press: Boston, MA, 1981.

(22) Stout, S. A.; Magar, V. S.; Uhler, R. M.; Ickes, J.; Abbott, J.; Brenner, R. Characterization of naturally occurring and anthropogenic PAHs in urban sediments—Wycoff/Eagle Harbor superfund site. *Environ. Forensics* **2001**, *2*, 287–300.

(23) Burns, W. A.; Mankiewicz, P. J.; Bence, A. E.; Page, D. S.; Parker, K. R. *Environ. Toxicol. Chem.* **1997**, *16*, 1119–1131.

(24) Neff, J. M. *Bioaccumulation in Marine Organisms. Effect of Contaminants from Oil Well Produced Water*; Elsevier: Boston, MA, 2002; pp 1–452.

(25) Short, J. W.; Rice, S. D.; Lindeberg, M. R.; Carls, M. G. Use of polyethylene membrane devices for monitoring diesel oil contamination on and within beaches. In Program and Abstracts, 31st Annual Aquatic Toxicity Workshop, Confederation of Aquatic Sciences, Charlestown, Prince Edward Island, Canada, October 24–27, 2004; p 25.

(26) Springman, K. Center for Health and the Environment, University of California, Davis, CA. Personal communication, 2005.

(27) Carls, M. G. Unpublished data, 2005.