8 Advantages of Quantitative Chemical Fingerprinting in Oil Spill Source Identification

Gregory S. Douglas, Scott A. Stout, Allen D. Uhler, Kevin J. McCarthy, and Stephen D. Emsbo-Mattingly

NewFields Environmental Forensics Practice LLC, 100 Ledgewood Place, Suite 302, Rockland, MA 02370

8.1 Introduction

The world economy is dependent on the exploration, production, transportation, and refining of petroleum. The demand for petroleum continues to increase as developing countries move from agricultural- to industrial-based economies. To meet the increased demand for oil over the next two decades, the world oil supply of crude oil will have to increase by over 45 million barrels/day (NRC, 2003), and it is apparent that most of this petroleum will be transported by large tankers (NRC, 2003). Much of the national and international shipping of raw materials and commercial goods in this growing world economy will also be transported by marine vessels. Thus, the risk of accidental discharges of petroleum to the environment — either as crude oil from production operations or as marine fuels from the discharge of maritime cargo vessels — will most certainly increase.

Oil spills can take place in rivers, open waters, and navigable coastal waterways. The natural resources damages (NRD) and liability associated with the release of even a small volume of petroleum warrants a thorough study of the fate of the spilled petroleum. The fundamental elements of almost all oil spill studies includes sufficient chemical characterization so investigators may (1) defensibly determine the source of the oil, (2) distinguish spilled oil from pre-existing background hydrocarbons, (3) quantifiably evaluate the extent of impacted ecosystems (Stout et al., 2001), and (4) reliably monitor oil spill cleanup and remediation.

Detailed chemical analysis of petroleum — often referred to as "chemical fingerprinting" — has played an important role in the identification of oil arising from accidental spills (Wang and Fingas, 1999; Wang et al., 2001; Stout et al., 2001; Page et al., 1995). The results of "chemical fingerprinting" are often supplemented by other lines of evidence (e.g., spill records, operational records, proximity of candidate sources to a slick, and oil slick trajectory modeling) (Lehr et al., 1999) in order to develop a comprehensive conceptual model of a spill event.

The modern chemical fingerprinting analytical methods used today have evolved over the last two decades, largely due to the development and increased sophistication of analytical instrumentation (Boehm et al., 1997; see also Chapter 1 herein). The cornerstone of modern petroleum fingerprinting is high-resolution capillary gas chromatography (GC). The GC technique provides a means to physically separate a complex mixture of hydrocarbons into the individual chemical compounds that can then be detected, identified, and measured by various means. The most appropriate and common means for the detection, identification, and measurement of the individual compounds comprising petroleum include GC in

combination with flame ionization detection (FID) or mass spectrometry (MS). Using both of these methods the relative abundance of an individual compound is converted to an electronic signal that is reflected as a peak on the resulting chromatograms. The magnitude of the peak represents the concentration of that compound in the mixture (Douglas et al., 1994). The exponential increase of low-cost computing power and data storage has also provided oil spill investigators with powerful statistical, numerical, and graphical analysis tools that can be used for quantitative chemical fingerprinting. The ability to store large amounts of raw instrument file outputs for reference oils and field samples has also accelerated the development of petroleum product reference libraries rich in information that can be easily accessed and evaluated as new methods and interpretive approaches are developed.

Petroleum is comprised of thousands of individual compounds, many of which can be separated by GC and measured by FID or MS. Thus, GC analysis provides a means of separating hydrocarbon chemicals, producing a distribution of peaks or varying proportions that represent the "fingerprint" of the oil. The GC "fingerprint" of the spilled oil can be compared to the fingerprints for any number of candidate source oils analyzed by the same method(s).

GC fingerprinting of spilled oils, candidate sources, and potentially impacted samples can be conducted qualitatively or quantitatively (see ahead). In turn, any correlations between a spilled oil and its potential sources or potentially impacted samples also can be made qualitatively or quantitatively. Stout et al. (2005) identified the strengths and weaknesses of each approach and concluded that quantitative correlations provide the most reliable and unbiased basis upon which the source or impact of spilled oil can be defensibly determined. One particular advantage of quantitative data is the ability to address the issue of spilled petroleum that is comprised of mixtures, such as (1) mixtures from the commingling of multiple spilled oils from the same or different sources or (2) mixtures of a spilled oil and any pre-existing oil (or other hydrocarbons) in the environment (Stout et al., 2005). If unrecognized, these complications can confound "spill oil-to-source oil" or "spill oil-to-impacted sample" correlations using some correlation statistical methods, such as the revised Nordtest approach (Daling et al., 2002) or the emerging European Committee for Standardization (CEN) protocol (Chapter 7). Mixing of this sort is not uncommon. For example, large marine vessels carry hundreds of thousands of gallons of fuels (diesel marine oil [DMO], intermediate fuel oil [IFO], heavy fuel oil [HFO]), which may be mixed internally within bilges or during fuel blending, or externally after a release due to physical mixing in coastal waters, ports, and harbors. Also, hydrocarbons from anthropogenic and natural sources are ubiquitous in the environment and, therefore, will be available to mix with any spilled oil. In this chapter we discuss the development and field validation of a quantitative chemical fingerprinting mixing model approach that will improve the resolution and accuracy of statistical correlations when mixing — either among multiple spilled oils or between a spilled oil and other, often pre-existing hydrocarbons in the environment — has potentially affected the samples under investigation.

8.2 Qualitative Fingerprinting Methods

The utility of qualitative fingerprinting methods such as ASTM 3328, ASTM 5739, and Nordtest (1991) in oil spill investigations have been previously discussed (Stout et al. [2005]; see also Chapter 1 herein). Qualitative chemical fingerprinting analysis of spilled oil, candidate sources, and background materials can be best described as a visual comparison between various spectroscopic or chromatographic fingerprints. Such comparisons inescapably introduce a degree of subjectivity to the source identification evaluation, which is an undesirable feature of science.

The qualitative approach to chemical fingerprinting of oil spills was formalized in two

standards of the American Society for Testing and Materials (ASTM), which are still used today in oil spill investigations in the U.S. ASTM D3328 is a qualitative GC/FID fingerprinting method (which can also include GC-flame photometric detection) that was originally approved in 1990 (ASTM, 1990). This method states,

> The matching of oil samples is essentially a profiling technique based on the premise that identical oils give identical chromatograms. Normally, the matching of a spilled oil to a suspect oil can be accomplished by comparison of the chromatograms for each of the oils in a spill case.

ASTM D3328 goes on to state that after considering the effects of weathering,

> Normally, a direct comparison of chromatograms will suffice for establishing identity or nonidentity between samples [and] if the chromatograms are the same on the basis of the peak-for-peak matching, there is a high degree of probability that the samples are from the same source.

This protocol defines the match criteria as follows:

Match — like the sample submitted for comparison, that is, the chromatographic pattern is a virtual overlay.

Probable match — the chromatographic pattern is similar to that of the samples submitted for comparison, except (a) for changes that could be attributed to weathering or (b) differences attributable to specific contamination.

Indeterminate — the chromatographic pattern is somewhat similar to that of the sample submitted for comparison, except for certain differences (due to weathering) that it make impossible to ascertain whether the unknown is the same oil heavily weathered, or a totally different oil.

Nonmatch — unlike the samples submitted for comparison.

These criteria introduce considerable subjectivity into the correlation analysis, the degree to which can vary depending upon the data quality and experience of the interpreter. The issue of data quality is very important since ASTM D3328 acknowledges that "no statement is made about either the precision of bias of this test method since the result merely states whether there is conformance to the criteria for comparison specified in the procedure." Ironically, strict interpretation of the "match" criteria (above) would mean that any weathering of the spilled oil, a phenomenon that is virtually guaranteed to occur when oil is spilled onto water, would prohibit ever concluding that any unweathered candidate source can be a "match" for a weathered spilled oil. This method, therefore, provides a relatively weak scientific basis upon which to assign multimillions of dollars of liability for spilled oil.

In 2000, ASTM introduced D5739, an additional procedure that relies upon GC/MS fingerprinting [American Society for Testing and Materials (ASTM), 2000]. The introduction of this procedure tends to acknowledge and substantiate the shortcomings of ASTM D3328. D5739 uses GC/MS to acquire a greater number of more detailed "fingerprints" for a prescribed list of compound groups found in any given oil sample, which undoubtedly provide a more detailed basis for comparing oil spills and their candidate sources. However, like ASTM D3328 (above), ASTM D5739 also relies upon interpretation of these fingerprints as "qualitative comparisons" and "direct visual comparison." The method concludes by stating the procedure is "a means of making qualitative comparisons between petroleum samples; quantitation of the various chemical components is not addressed." Thus, despite acquiring chemical fingerprinting data of increasing specificity using GC/MS, the method's reliance upon "direct visual comparison" of the resulting fingerprints by "placing the EIC's (fingerprints) one over the other," again introduces a significant degree of subjectivity into the source identification process. This method defines its own match criteria, summarized as follows:

Similar — all the fingerprints examined for two oil samples display the same qualitative

pattern by visual examination . . . except those attributable to the precision of the analysis (as determined by analysis of a sample in duplicate) or weathering.

Inconclusive — all the fingerprints for two oil samples display no qualitative differences except one or more slightly greater than the precision of the analysis (as determined by analysis of a sample in duplicate). In this case the questionable oil should be reprepared and re-analyzed if possible and be deemed inconclusive if similar results are obtained.

Dissimilar — some or all of the fingerprints for two oil samples display discrepancies, particularly among the most weathering-resistant target compounds.

These criteria are highly subjective and prone to interpretation differences. Furthermore, since two of the match criteria are based on the sample duplicate precision with no stated minimum performance criteria, it is possible that one laboratory with poor precision will achieve more "matches" than another laboratory with excellent precision. Therefore, the absence of quality control renders this method poorly suited for the definitive determination of liability or guilt in the case of mystery oil spills.

In summary, qualitative chemical "fingerprinting" analysis of spilled oil, candidate sources, and background materials can best be described as a visual comparison between various chromatographic or other forms of chemical fingerprints. Such comparisons inescapably introduce a degree of subjectivity that can be undesirable when subtle differences, environmental weathering, data from multiple sources or vintages, and mixed petroleums are significant factors in the evaluation.

8.2.1 Shortcomings of Qualitative Fingerprinting

For some oil spill investigations, the qualitative approaches of ASTM D3328 and D5739 or Nordtest (1991) may be sufficient to reach defensible conclusions regarding the source of an oil spill. These investigations are generally limited to situations in which there are markedly disparate candidate source oils, only one of which can reasonably match the spilled oil and where extensive weathering of the sample has not occurred. However, most oil spill studies are not this straightforward. At least four particular circumstances in which qualitative fingerprinting can be problematic are spill situations involving (1) significant weathering of the spilled oil, (2) comparison of genetically similar spill and source oils or fuels, (3) qualitatively similar spill and source oils but with varying concentrations, and (4) mixing of spill oils with each other or with preexisting hydrocarbons in the environment. These situations are briefly described ahead.

8.2.1.1 Weathered Oils

The first of these shortcomings, weathering, is acknowledged by the ASTM D3328 method's match definitions described previously. Specifically, fingerprinting differences attributable to weathering require ASTM D3328 to conclude, at best, a *probable match* exists. Even if qualitative matches are achieved between a spill and source oil using ASTM D5739, the strongest conclusion one may reach is that the oils are "similar." These ASTM-defined conclusions, "probable match" or "similar" are not particularly useful conclusions for litigious situations. Thus, weathering is an inevitable shortcoming in the application of the qualitative ASTM methods or of Nordtest (1991).

8.2.1.2 Genetically Similar Oils

It is common for oils produced, transported, and refined within a given geographic province to share a certain degree of chemical similarity. This stems from the fact that crude oil generated within a specific geologic province will, as a consequence of similar character of the ancient organic matter and similar subsurface heating conditions that produced the crude (Tissot and Welte, 1984), yield crude oils of a related chemical character. Very subtle differences between genetically similar oils from a

given oil province, field, or even well, can still be recognized with the appropriate chemical data (e.g., Kaufman et al., 1990; McCaffrey et al., 1996; Nicolle et al., 1997; Hwang et al., 2000). When these crude oils are subsequently refined into specific petroleum products at nearby refineries (e.g., IFO), the resulting petroleum products often will also exhibit genetically similar chemical features "inherited" from the parent crude oil (Peters et al., 1992; see also Chapter 1 herein). Furthermore, as these petroleum products are then distributed throughout the market, multiple handlers (e.g., terminal, pipelines, and tankers) may carry the same fuel. The difference between the characteristics of a fuel used on one vessel versus that used by another may only depend upon the nature of any previous fuel that was present in their tanks at the time of refueling.

In summary, the chemical differences among crude oils or fuels in a geographic area may be very subtle, and therefore go unrecognized using qualitative fingerprinting techniques. This particular shortcoming was the basis for revision of the Nordtest (1991) protocol that was commonly applied to oil spills in the North Sea, where genetically similar crude oils and fuels co-occur, thereby confounding too many oil spill investigations in which only qualitative fingerprinting data were available (Daling, personal communication, 2000).

8.2.1.3 Qualitatively Similar Oils

It is conceivable that some oils may exhibit qualitatively similar fingerprints but differ in the absolute concentrations of different compounds. Consider the situation where the patterns of pentacyclic triterpanes and steranes in two oils are qualitatively similar, but the concentration of triterpanes is much higher in one oil than in the other. This type of difference would go unnoticed in a qualitative fingerprinting protocol such as ASTM D5739 or Nordtest (1991). While this type of situation may seem extraordinary — in part, perhaps because it is not commonly pursued — it can occur. An example from a mystery spill investigation is provided here.

In this case, another laboratory analyzed two oils (a spill and presumed source) via ASTM D3328 and D5730 protocol. A qualitative interpretation of these data concluded that the two oils were a "probable match" in accordance with ASTM D3328 and D5739. Some justification for this conclusion is evident upon inspection of the qualitatively similar GC/FIDs and selected extracted ion plots (EIPs) shown in Figure 8-1. (These data, though representative of the data from the other laboratory, were generated in our laboratory using the quantitative analysis as described in Section 8.3.) The other laboratory had considered the minor differences observed in the GC/FIDs (e.g., disparate UCM profiles, isoprenoid distributions, and Pr/Ph ratios) as being attributable to weathering. Similarly, their qualitative comparisons among the various EIPs obtained via ASTM D5739 concluded the two oils were correlated. While indeed the oils are qualitatively similar, minor differences can be observed (e.g., relative abundance of oleanane). During a qualitative comparison such minor differences might go unnoticed. This is an obvious shortcoming of qualitative interpretations. Of particular importance in this study, however, were the qualitatively similar distributions of C2-dibenzothiophenes (D2) and C2-phenanthrenes (P2) in which both oils showed almost identical *m/z* 212 and *m/z* 206 patterns, respectively (Figure 8-1). However, quantitative analysis of the absolute concentrations of the D2 and P2, as well as the D3 and P3, indicated that the two oils contained disparate absolute concentrations (Figure 8-2). This important difference could not be observed upon qualitative comparison of the EIPs patterns using ASTM D5739. Consequently, we concluded that the two oils, though qualitatively similar, were definitively not the same oil.

8.2.1.4 Mixing

In addition to changes brought about by weathering, the chemical fingerprint for a spilled oil will change if it becomes mixed with other oils released concurrently into the environment or

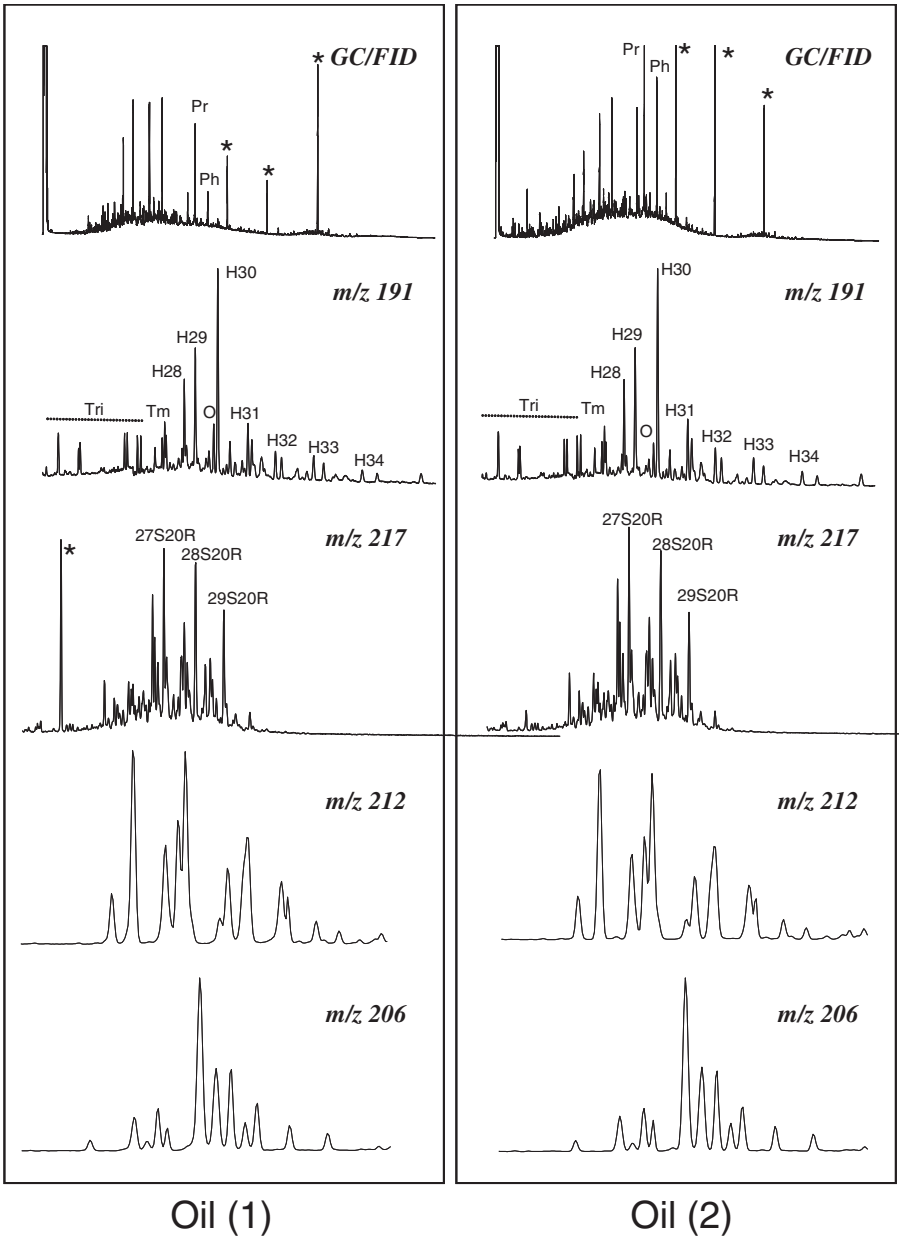

Figure 8-1 Comparison of two oils that exhibit qualitatively similar GC/FID patterns and selected extracted ion profiles (partial) that were interpreted to be "probable matches" by a laboratory conducting ASTM D3328 and D5739. Quantitative analysis showed some measurable differences existed, including higher concentrations of C2- and C3-dibenzothiophenes in oil 2 (see Fig. 8-2). *—internal standards.

with any pre-existing oil (or other hydrocarbons) already present in the environment. For example, large merchant vessels may store hundreds of thousands of gallons of diesel, intermediate, and heavy fuels, which, if released during an accident or illegal discharge, may yield variably "mixed" fingerprints that would not qualitatively match any of the individual bunkered fuels. Qualitative comparison of discrete source samples to the

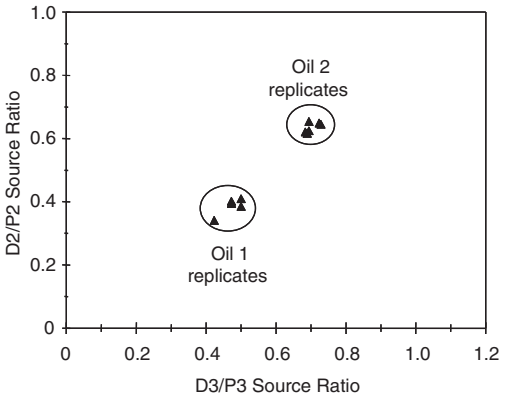

Figure 8-2 Double-ratio plot showing the results of replicate analysis of the two oils shown in Figure 8-1. Despite qualitatively comparable D2 (*m/z* 212), D3, P2 (*m/z* 206), and P3 patterns (e.g., Fig. 8-1) the concentration of D2 and D3 were higher in Oil 2.

spill samples, which are composed of variable mixtures of cargo and/or fuel, would fall short of recognizing the true nature of such a mixture. Similarly, if oil is spilled into an area already containing pre-existing "background" oil or other hydrocarbons, the resulting fingerprint will be qualitatively different from the source oil. Thus, under conditions of potential mixing, qualitative fingerprinting can be easily confounded and result in false negative correlations (e.g., fingerprint identification is a *nonmatch* or is *dissimilar*).

8.3 Quantitative Fingerprinting Methods

In order to confidently identify the source(s) of weathered, genetically, or qualitatively similar oils, or oil mixtures in environmental samples, some form of quantitative fingerprinting is more likely to yield more defensible results (Stout et al., 2005). The principal reason for this is that recognizing the often subtle differences among spill and source oils cannot reasonably rely upon the qualitative ASTM methods just described. As a result, over the past 15 years there has been an acknowledged need for and resulting advances toward quantification in the chemical fingerprinting of oil spill investigations.

8.3.1 Semiquantitative versus Fully Quantitative Methods

"Quantification" can mean different things to different laboratories. In some environmental laboratories quantitative data refers to the determination of peak areas (or sometimes heights) of compounds within a single chromatogram or extracted ion profile. These data are more accurately considered *semiquantitative* as they do not rely upon the use of internal or external standards or response factors, nor do they result in determination of concentrations of individual target compounds or compound groups. In addition, the quality of peak area or peak height data is frequently not analyzed with a multiconcentration initial calibration that demonstrates the applicable linear range of the target peaks. Continuing calibration standards are also frequently absent, which results in the failure to demonstrate the stability of the instrument over time. Nevertheless, semiquantitative data are still more useful than qualitative data (the latter of which rely entirely upon visual examination of chromatographic patterns). For example, diagnostic ratios can be developed from peak areas and allow statistical comparison between samples, as is described in Chapter 7. However, the use of diagnostic ratios generated from semiquantitative peak height or area responses still cannot address the issue of mixing between oils or between spilled oils and pre-existing hydrocarbons. The reason is that ratios do not mix linearly. A simple example of this is shown in the following data.

	Oil #1	*Oil #2*	*50 : 50 Mix*
Peak A (Response)	1,000	1,000	1,000
Peak B (Response)	5,000	10,000	7,500
Ratio A/B	0.20	0.10	0.13

In this simple example, assume that 1 µL of two oils has been injected and the responses for two peaks (A and B) were measured. The ratios of A/B in oil #1 and #2 are 0.2 and 0.1, respectively. Because the responses (a reflection of their relative concentrations) of peaks A and B vary between the oils, a 50 : 50 mix of

these oils yields a ratio of 0.13 and not 0.15, which might be anticipated if the original A/B ratios (0.20 and 0.10) in each oil were mathematically mixed 50:50. It is possible for semiquantitative peak areas or heights to be normalized to the mass of oil injected into the gas chromatograph — and all chromatographic conditions are kept constant (e.g., split ratios). Under these conditions, the measured responses (but not the ratios) could be used in addressing the issue of mixing.

These potential difficulties are avoided if internal standards and multilevel calibration standards, the latter containing a suite of structurally similar representatives of the targeted compounds, are employed to determine the absolute concentrations of target analytes in the samples studied. This approach is considered fully *quantitative* and is described further ahead.

The investigations following the *Exxon Valdez* oil spill in 1989 first promulgated the use of fully quantitative fingerprinting methods in oil spill investigations (Douglas et al., 1996; Kvenvolden et al., 1995; Page et al., 1995; Bence and Burns, 1995; Boehm et al., 1997). In these studies the absolute concentrations of the target analytes included those from ASTM D5739, as well as other diagnostic analytes derived from the petroleum geochemistry literature (e.g., Peters and Moldowan, 1993; Peters et al., 2005a, 2005b; Radke et al., 1986). These chemical concentrations in oils, sediments, and tissues were measured quantitatively, thereby providing absolute concentrations of diagnostic chemicals (Page et al., 1999; Stout et al., 2001; Daling et al., 2002a; *Federal Register*, 1994). Thus, instead of a qualitative, visual comparison of chromatographic peaks or semiquantitative comparisons of ratios based upon peak heights, tables of the numerical concentration data were developed and used for the calculation of diagnostic ratios (Stout et al., 2001; Page et al., 1995; Douglas et al., 1996; Douglas and Uhler, 1993) and other parameters. These factors allow simultaneous, statistical, or numerical comparisons among many samples and even allocation among disparate sources (Burns et al., 1997). This quantitative approach formed the basis for modern oil spill fingerprinting methods (Boehm et al., 1997).

Briefly, fully quantitative fingerprinting requires that the concentrations of individual and homologue series of diagnostic or environmentally important hydrocarbons be determined using the internal standard method. In this approach, recovery and quantitation internal standards, typically deuterated analogues of PAHs or *n*-alkanes, are added to the samples in known concentrations prior to GC/FID and GC/MS analysis. The internal standards serve as internal references in quantification of target compounds and as surrogates to judge method performance. Overarching the analytical methods are rigorous quality-control procedures that ensure accuracy and precision, and minimize bias. Typical QC protocols include procedural blanks, replicates, laboratory control samples, and reference oils (Page et al., 1995; Douglas et al., 1994; *Federal Register*, 1994).

Use of these techniques produces precise and accurate numerical concentration data that allow for comparison of spilled oil to candidate source oils using graphical (Boehm et al., 1997), statistical (Daling et al., 2002a; Christensen et al., 2004), or numerical analysis tools (Urdal et al., 1986; Burns et al., 1997; Stout et al., 2001; Mudge, 2002). This type of comparison reduces the subjectivity of the interpreter with greater reliance on data objectivity. These mathematically based approaches to fingerprinting interpretation also permit intercomparison of many oils to each other simultaneously, thereby improving the investigator's ability to identify similarities and/or differences among impacted field samples and suspect sources. The quantitative fingerprinting approach can identify the often subtle chemical relationships that help correlate or decouple samples and candidate sources. It is these subtle, but very important, features that often go unrecognized in qualitative fingerprinting methods (ASTM D3328 and 5739) in which patterns are compared visually.

Concentrations of polycyclic aromatic hydrocarbons (PAHs) and petroleum biomarkers

(or biological markers) are particularly useful quantitative measures that benefit oil spill investigations because of their characteristic distributions and environmental recalcitrance (Douglas et al., 1996; Wang et al., 1999; Stout et al., 2001, 2000). Biomarkers are naturally occurring, ubiquitous, and stable hydrocarbons that occur in crude oils and most petroleum products (Peters and Moldowan, 1993). The distribution of biomarkers in oil reflects the "genetic" history of that oil, specifically, the type(s) of precursor ancient organic matter and the thermal history of the rock strata in which the oil was formed. Because of the diversity of organic matter and oil-forming environments, there are varied biomarker patterns among oils found in the many petroleum reservoirs around the world (Peters et al., 1993; see also Chapter 3).

Quantitative PAH and biomarker data form the basis of the oil spill investigation protocols and have been used in many oil spill investigations around the world (MSRC, 1995; Wang and Fingas, 1999; Wang et al., 1999). In virtually all the applications of quantitative oil spill investigations, the data have been evaluated using some kind of graphical or advanced data analysis tool (Urdal et al., 1986; Page et al., 1995; Henry et al., 1997; Burns et al., 1997; Stout et al., 2000, 2001; Mudge, 2002). In the sections that follow, we present an overview of the methods used in the collection of quantitative chemical fingerprinting data and the protocols by which these quantitative data can be applied to oil spill studies.

8.3.2 Data Generation for Fully Quantitative Fingerprinting

8.3.2.1 Sample Collection

The success of any oil spill investigation begins with collection of representative samples, using appropriate techniques (see Chapter 2). Protocols for the collection of the floating oil, mousse, tarballs, sheens, vessel tanks, and port supplier tanks have been described adequately (e.g., Nordtest, 1991). In addition, some natural resource damage assessment (NRDA) guidelines have adequately described sampling of oiled shorelines, sediments, animals, and vegetation (e.g., PERF, 1995).

In the immediate aftermath of an oil spill, it is crucial to identify and sample all the viable candidate sources. Identification of appropriate candidate sources is accomplished by a combination of common sense (e.g., collection of obvious source samples), combined with a thorough record review (e.g., sailing and cargo records, fueling records, etc.) for identification of important, but less obvious, sources. In many cases, oil spill samples are collected by personnel from the investigating agency who often provide split samples to potentially responsible party (PRP) representatives. Ideally, the PRP representative should collect samples of fugitive and source oils in parallel with regulatory investigators.

The typical types of samples collected in a harbor/port spill setting include samples of floating fugitive oil, oil stranded on shorelines and pilings, and oiled animals or water birds. Candidate source samples include representatives of ship cargo and fuel oil (e.g., fuel tank oils, bilge tank oils, oil water separator, settling tank oils, post oil sensor lines, external hull wipes, and drain tank oils from suspect vessels), as well as oil from any nearby port's fuel storage facility. In coastal investigations, oil/sheens near storm drain outfalls should also be collected if present. Often, oil on water is observed as sheen — an ultrathin accumulation of oil floating on water. Because of their nature, sheen samples cannot be collected using simple grab sample techniques. Rather, sheen samples are collected using a 4-inch-diameter, highly porous, TFE-fluorocarbon polymer net. This net was developed in conjunction with the U.S. Coast Guard to capture thin oil sheens in surface water (General Oceanics, 2006). The net is passed through the sheen five times, removed from the net holder ring, and placed into a glass jar for shipment. In our experience, we have found that the Teflon net must be rigorously precleaned with methylene chloride and blank tested prior to use. The precleaned nets should be stored in a clean environment and used within 1–3

months of cleaning. In all cases, appropriate blanks must accompany the samples to the laboratory (e.g., wipe sample and sheen blanks). Samples of free oil, sheen netting, sediments, and other oiled media should be stored at 4°C immediately after collection. Storage of oil samples at ambient temperatures in sealed glass containers will, in most cases, not impact the integrity of the sample for forensic studies.

8.3.2.2 Sample Preparation

Spilled oil/tarball and candidate source oil samples are prepared for analysis by weighing approximately 50 mg of oil into a tared 10-mL volumetric flask, bringing it to volume in dichloromethane (DCM) and removing a 1-mL aliquot for analysis. If a recovered oil sample is suspected to contain entrained water, the water in the extract is removed with anhydrous sodium sulfate prior to taking a 1-mL analytical aliquot. Each analytical batch of authentic samples ($n < 20$) should include appropriate quality-control samples, for example, a procedural blank (PB: 1 ml of DCM), a laboratory control sample (LCS) consisting of 1 mL of DCM spiked with selected hydrocarbons in known concentrations to monitor method accuracy, and one set of triplicate oils (i.e., a single oil prepared three times by drawing separate aliquots and spiking each individually) as a measure of precision and reproducibility of the data. After spiking the 1-ml aliquots of each sample with surrogate internal standard (SIS; *o*-terphenyl, naphthalene-d_8, phenanthrene-d_{10}, chrysene-d_{10}, or equivalent) and recovery internal standard (RIS; 5α-androstane, acenaphthene-d_{10}, fluorene-d_{10}, benzo[a]pyrene-d_{12}, or equivalent), the samples are split for quantitative GC/FID and GC/MS analysis.

Sheen samples are removed from the sample jar with precleaned tweezers and placed in a widemouth 250-mL jar with a Teflon lined cap for extraction. The sample is spiked with surrogate compounds and serially extracted two times with 50 mL of methylene chloride. The sample container should be rinsed with the first 50 mL of extraction solvent. The combined extract is concentrated to 1 mL by Kuderna Danish/nitrogen evaporation methods, spiked with internal standards, and analyzed.

8.3.2.3 GC/FID Analysis

A high-quality GC/FID analysis provides a hydrocarbon "fingerprint" of the sample, an assessment of the degree of sample weathering, and quantitative measures of important compositional and source-specific chemical compounds. Target analytes commonly measured by GC/FID analysis are listed in Table 8-1. The measured parameters include the total petroleum hydrocarbon (TPH) carbon ranges of *n*-C_{10} to *n*-C_{44+}, and *n*-C_{10} to *n*-C_{28} (diesel range organics [DRO]) and a variety of individual target hydrocarbons including *n*-alkanes and isoprenoid compounds.

Gas chromatography-flame ionization detection (GC/FID) should be conducted using a modern capillary gas chromatograph with a splitless injection port. In our work, we use an Agilent 6890 GC. The gas chromatograph should be fitted with a 60 m × 0.32 mm ID, 0.25 µm film thickness, DB-5 capillary column (or equivalent). A suitable GC oven program should be used to facilitate baseline resolution of *n*-alkane and isoprenoid hydrocarbons (e.g., *n*-C_{17} versus pristane); the program we favor begins with an initial temperature of 40°C (1 min), followed by a 6°C/min ramp rate to a final temperature of 315°C, followed by a 30-min hold. Ideally, hydrogen should be used as the carrier gas.

Prior to sample analysis, a minimum five-point calibration is performed to demonstrate the linear range of the analysis. The calibration solution is composed of selected aliphatic hydrocarbons within the *n*-C_9 to *n*-C_{40} range. Analyte concentrations in the standard solutions range from 1 ng/µL to 200 ng/µL. Target analytes not in the calibration solution may be quantified with the average relative response factor (RRF) of the nearest eluting compound(s). In our work, we assign the following response factors to certain target compounds: RRF of *n*-C_{14} assigned to 1380 and 1470 (C_{15} isoprenoids); RRF of *n*-C_{16} is

Table 8-1 Inventory of Target Saturated Hydrocarbons (SHC), PAHs, and Biomarkers Commonly Analyzed by GC/FID and GC/MS-SIM in Oil Spill Studies

Target SHC and PAH	*Key*	*Quant. Ion (m/z)*	*RF*	*Target Biomarkers*	*Key*	*Quant. Ion (m/z)*	*RF*
SHC — GC/FID		NA	RF_{alk}	**Sesquiterpanes**	SQx	123	Decalin
n-C_8 > *n*-C_{40}, pristane, phytane	*n*-Cx	NA	RF_{alk}	**Alkylcyclohexanes**	CH-X	83	Alk
nor-pristane, TPH (C8–C44)		NA	RF_{alk}				
PAH Groups — GC/MS		NA		**Tricyclic Triterpanes — GC/MS**			
Naphthalene	N	128	N	C_{23} Tricyclic triterpane	TC23	191	Hop
C1-Naphthalenes	N1	142	N	C_{24} Tetracyclic triterpane	TC24	191	Hop
C2-Naphthalenes	N2	156	N	C_{25} Tricyclic triterpane	TC25	191	Hop
C3-Naphthalenes	N3	170	N	C_{26} Tricyclic triterpanes	TC26	191	Hop
C4-Naphthalenes	N4	184	N	C_{28} Tricyclic triterpanes	TC28	191	Hop
Acenaphthene	ACE	154	ACE	C_{29} Tricyclic triterpanes	TC29	191	Hop
Acenaphthylene	ACY	152	ACY	C_{30} Tricyclic triterpanes	TC30	191	Hop
Biphenyl	BPHN	154	BPHN				Hop
Dibenzofuran	DBF	168	DBF	**Pentacyclic Triterpanes — GC/MS**			Hop
Fluorene	F0	166	F	18α(H)-22,29,30-trisnorhopane (**T_s**)	27Ts	191	Hop
C1-Fluorenes	F1	180	F	17α(H)-22,29,30-trisnorhopane (**T_m**)	27Tm	191	Hop
C2-Fluorenes	F2	194	F	17α(H),21β(H)-28,30-bisnorhopane	28ab	191	Hop
C3-Fluorenes	F3	208	F	17α(H),21β(H)-25-norhopane	25nor	191	Hop
Dibenzothiophene	DBT	184	D	17α(H),21β(H)-30-norhopane	29ab	191	Hop
C1-Dibenzothiophenes	D1	198	D	18α(H)-30-norneohopane ($C_{29}T_s$)	29Ts	191	Hop
C2-Dibenzothiophenes	D2	212	D	17β(H),21α(H)-normoretane	29ba	191	Hop
C3-Dibenzothiophenes	D3	226	D	18α(H) and 18β(H) oleanane	30O	191	Hop
C4-Dibenzothiophenes	D4	240	D	17α(H),21β(H)-hopane (**Hop**)	30ab	191	Hop
Phenanthrene	P0	178	P	17β(H),21α(H)-moretane	30ba	191	Hop
Anthracene	A0	178	A	22S-17α(H),21β(H)-30-homohopane	31abS	191	Hop
C1-Phenanthrenes/ Anthracenes	P1 or PA1	192	P	22R-17α(H),21β(H)-30-homohopane	31abR	191	Hop
C2-Phenanthrenes/ Anthracenes	P2 or PA2	206	P	Gammacerane	30G	191	Hop
C3-Phenanthrenes/ Anthracenes	P3 or PA3	220	P	22S-17α(H),21β(H)-30-bishomohopane	32abS	191	Hop
C4-Phenanthrenes/ Anthracenes	P4 or PA4	234	P	22R-17α(H),21β(H)-30-bishomohopane	32abR	191	Hop
Fluoranthene	FL	202	FL	22S-17α(H),21β(H)-30-trishomohopane	33abS	191	Hop
Pyrene	PY	202	PY	22R-17α(H),21β(H)-30-trishomohopane	33abR	191	Hop
C1-Fluoranthenes/Pyrenes	FP1	216	FL	22S-17α(H),21β(H)-30-tetrakishomohopane	34abS	191	Hop
C2-Fluoranthenes/Pyrenes	FP2	230	FL	22R-17α(H),21β(H)-30-tetrakishomohopane	34abR	191	Hop
C3-Fluoranthenes/Pyrenes	FP3	244	FL	22S-17α(H),21β(H)-30-pentakishomohopane	35abS	191	Hop
C4-Fluoranthenes/Pyrenes	FP4	258	FL	22R-17α(H),21β(H)-30-pentakishomohopane	35abR	191	Hop
Naphthobenzothiophenes	NBT	234	NBT	**Steranes — GC/MS**			
C1-Naphthobenzothiophenes	NBT1	248	NBT	13β,17α-diacholestane(20S)	27dbS	217	Chol
C2-Naphthobenzothiophenes	NBT2	262	NBT	13β,17α-diacholestane(20R)	27dbR	217	Chol
C3-Naphthobenzothiophenes	NBT3	276	NBT	5α,14β,17β-cholestane(20R)	27bbR	218	Chol
C4-Naphthobenzothiophenes	NBT4	290	NBT	5α,14β,17β-cholestane(20S)	27bbS	218	Chol
Benzo[a]anthracene	BAA	228	BAA	5α,14α,17α-cholestane(20R)	27aaR	217	Chol
Chrysene	C	228	C	5α,14β,17β,24-methylcholestane(20R)	28bbR	218	Chol
C1-Chrysenes	C1	242	C	5α,14β,17β,24-methylcholestane(20S)	28bbS	218	Chol
C2-Chrysenes	C2	256	C	5α,14α,17α,24-methylcholestane(20R)	28aaR	217	Chol
C3-Chrysenes	C3	270	C	5α,14α,17α,24-ethylcholestane(20S)	29aaS	217	Chol
C4-Chrysenes	C4	284	C	5α,14β,17β,24-ethylcholestane(20R)	29bbR	218	Chol
Benzo[b]fluoranthene	BBF	252	BBF	5α,14β,17β,24-ethylcholestane(20S)	29bbS	218	Chol
Benzo[j,k]fluoranthene	BKJF	252	BKJF	5α,14α,17α,24-ethylcholestane(20R)	29aaR	217	Chol
Benzo[a]fluoranthene	BAF	252	BAF	5β(H)cholane (**Chol**)	Chol	217/218	
Benzo[a]pyrene	BAP	252	BAP	**Triaromatic Steroids (TAS) — GC/MS**			Chol
Benzo[e]pyrene	BEP	252	BEP	C_{20} — TAS	C20TA	231	Chol
Perylene	PER	252	PER	C_{21} — TAS	C21TA	231	Chol
Indeno[1,2,3-c,d]pyrene	IND	276	IND	C_{26},20S — TAS	SC26TA	231	Chol
Dibenzo[a,h]anthracene	DAH	278	DAH	C_{26}20R + C_{27},20S — TAS	RC26TA SC27TA	231	Chol
Benzo[g,h,i]perylene	BGHI	276	BGHI	C_{28},20S — TAS	SC28TA	231	Chol
Single PAHs				C_{27},20R — TAS	RC27TA	231	Chol
4-Methyldibenzothiophene	4MD	198	D	C_{28},20R — TAS	RC28TA	231	Chol
2/3-Methyldibenzothiophene	2MD	198	D				
1-Methyldibenzothiophene	1MD	198	D	**Decalins**			
3-Methylphenanthrene	3MP	192	P	Decalin	DE	138	DE
2/4-Methylphenanthrene	2MP	192	P	C1-Decalins	DE1	152	DE
2-Methylanthracene	2MA	192	P	C2-Decalins	DE2	166	DE
9-Methylphenanthrene	9MP	192	P	C3-Decalins	DE3	180	DE
1-Methylphenanthrene	1MP	192	P	C4-Decalins	DE4	194	DE
Retene	R	234	P				
Cadalene	CD	198	N				
5-Methylchrysene	MC	242	C				

assigned to 1650 (i.e., nor-pristane). Carbon range TPH analyses are performed by employing the baseline integration technique of Douglas et al. (1994). The "window" for each TPH range is determined from the *n*-alkane calibration standard. All calibration solution compounds that fall within the appropriate carbon range window are used to generate the average RRF for a given carbon range. Areas for surrogate and internal standard compounds are subtracted from the range integration prior to computation of the TPH.

Instrument calibration is assured by analysis of a mid-level calibration check standard after every 10 authentic samples. The check standard's response is compared versus the average RF of the respective analytes contained in the initial calibration. All authentic samples and quality-control samples are bracketed by passing mid-check standards.

8.3.2.4 GC/MS Analysis

Gas chromatography/mass spectrometry (GC/MS) analysis of oil spill samples arguably provides the most important quantitative "fingerprinting" data used in oil spill correlation analysis. We utilize an Agilent 6890 GC interfaced to a Hewlett-Packard 5973 mass selective detector (MSD). The GC should be equipped with a 60 m × 0.25 mm ID, 0.25 µm film thickness, DB-5 capillary column (or equivalent). The GC oven program we use begins at an initial temperature of 40°C with a 1-min hold, followed by a 6°C/min oven ramp to a final temperature of 315°C, followed by a 30-min hold time. Helium is the preferred carrier gas in this quadrapole mass spectrometry analysis.

The GC/MS is calibrated with perfluorotributylamine (PFTBA) at the beginning of each analytical sequence. A minimum 5-point initial calibration consisting of selected target compounds is established to demonstrate the linear range of the analysis (Table 8-1). Analyte concentrations in the standard solutions range from 0.01 ng/µL to 10 ng/µL.

Data acquisition is performed in the select ion monitoring (SIM) mode for optimal sensitivity and selectivity. Quantification of target compounds is performed by the method of internal standards using the average relative response factor of the parent compounds or otherwise representative compounds (Table 8-1) determined from the 5-point initial calibration (*Federal Register*, 1994).

Target compounds measured in this analysis include the polycyclic aromatic hydrocarbons (PAH) and the tricyclic and pentacyclic triterpanes and steranes listed in Table 8-1. Many alkyl PAH isomers are commercially unavailable as analytical standards. Thus, the quantitative response factors for such compounds are based upon that of their respective parent PAH (Table 8-1; *Federal Register*, 1994). Similarly, most target biomarkers are not contained in the initial calibration solution (largely due to the high costs of standards or lack of availability). The RFs for biomarkers are based on those of two representative compounds (Table 8-1). In the case of selected PAH groups, individual PAH isomers are also quantified using RFs derived for the individual isomers contained in the calibration solutions. Among these are four methyl-dibenzothiophene isomers, five methyl-phenanthrene isomers, retene (1-methyl-7-isopropyl-phenanthrene), cadalene (1,6-dimethyl-4-(1-methylethyl)-naphthalene), and 5-methylchrysene.

PAH and biomarker concentrations (as well as alkanes, described above) are calculated by the method of internal standards using the following:

$$C_a = [(A_a/A_i) \times (\mathrm{Amt}_i/\mathrm{RF}_i) \times D]/V_a$$

where

C_a = concentration of target analyte
A_a = area of quantification ion for target analyte
A_i = area of quantification ion for RIS
Amt_i = amount of RIS added to sample
RF_i = average RF for analyte determined from intial 5-point calibration
D = dilution factor (if applicable)
V_a = sample size (volume or mass)

Biomarker identifications are based upon comparison to selected authentic standards (Chiron

Laboratories), elution patterns in the peer-reviewed literature (e.g., Philp, 1985), and mass spectral interpretation from full-scan GC/MS analyses conducted in the laboratory. Triterpane concentrations are calculated using the RRF of C_{30} 17α(H), 21β(H)-hopane relative to the internal standard d_{12}-chrysene. Sterane concentrations are calculated using the RRF of 5β(H)-cholane relative to the internal standard d_{12}-chrysene.

8.3.2.5 Data Quality

Samples are analyzed in exclusive analytical batches, on the same GC/MS instrument (when possible) and, to the extent possible, within the same analytical sequence. Data analyses (peak integrations) are conducted by a single GC/MS analyst with experience in PAH and biomarker pattern recognition. Each of these steps, combined with the calibrations (initial and continuing), procedural blanks (PB), laboratory control samples (LCS), control oil analysis, surrogate recoveries, and duplicate/triplicate analyses, provides measures of data quality. Typical data quality objectives for these analyses are listed in Table 8-2.

8.3.3 Selection of Diagnostic Indices

Although the comparison of absolute concentrations of target compounds can be used to compare spill and source oils, more often than not diagnostic ratios calculated from these concentrations offer some advantages when comparing samples. One benefit of comparing the diagnostic indices rather than absolute concentrations of spilled oil and suspected source oils or potentially impacted samples is that the analytical variability due to concentration effects is minimized. This technique has been used in other oil spill investigation protocols in which semiquantitative data (e.g., peak heights or areas; Daling et al., 2002a, 2002b;

Table 8-2 Summary of Typical Data Quality Objectives Used in Quantitative Fingerprinting Studies of Waterborne Oil Spills

QC Element or Sample Type	*Minimum Frequency*	*Data Quality Objective/Acceptance Criteria*
MS Tuning	Prior to each run sequence using PFTBA	m/e 69: Base Peak (~100,000 counts minimum) m/e 219: 30–60% Base Peak abundance m/e 502: 2–8% Base Peak abundance
Initial Calibration	Prior to every instrument batch sequence or as needed indicated by continuing calibration check	5 point curve, minimum of 5 point. %RSD ≤25% for 90% of analytes and ≤35% for all analytes
Continuing Calibration	Must end analytical sequence, and every 12 field samples or 16 hours, whichever is more frequent	%RSD ≤25% for 90% of analytes. %RSD ≤35% for all analytes
Procedural Blank	Every batch/every 15–20 samples	Less than the reporting limit unless analyte not detected in associated sample(s) or associated sample analyte concentration is >5× blank value
Laboratory Control Sample (LCS)/LCS Duplicate	Every batch/every 15–20 samples	50%–130% recovery SVOCs
Recovery/Surrogate Standards	Every sample	50%–130% recovery for other SVOCs
Internal Standard (IS)	Every sample	50%–200% of the area of the IS in the associated calibration standard SVOC
Instrumental Check	One per initial calibration	80%–120% recovery
Control Oil	One per initial calibration One per sequence for FID	65%–135% recovery PAHs and hopane for biomarkers (for NSC only) alkanes — for chromatogram only

Faksness et al., 2002; Henry et al., 1997) or absolute concentration data are converted to useful quantitative ratios (Burns et al., 1997; Stout et al., 2001; Page et al., 1995). On the other hand, because each index depends on the responses or concentrations of multiple analytes, any errors in these may be increased in the calculation of a ratio. Thus, the precision of the indices is more important than the precision of the concentrations and must be evaluated prior to use in oil spill correlation analysis. It is our experience that in order for a ratio to be useful for interpretive purposes, the analytical precision of a representative oil sample replicate ($n = 3$) must generally be 10% (residual standard deviation) or less. The emerging CEN protocol described in Chapter 7 suggests the use of 14% RSD as a suggested measure of precision.

The data acquisition described earlier provides a set of GC/FID chromatograms, selected extracted ion profiles from GC/MS-SIM, and the tabulated absolute concentration data for each target analyte or analyte group (Table 8-1) identified in the field and QC samples. Evaluations of the biomarker, PAH, total petroleum hydrocarbon (TPH), and selected alkane concentration data are subsequently used to generate a suite of diagnostic indices allowing comparison among the spilled oil, candidate sources, and potentially impacted samples (Table 8-3).

The indices listed in Table 8-3 each serve some diagnostic purpose, but each is not necessarily useful in correlating all types of spilled oil to candidate sources, primarily due to the effects of weathering on some indices (Douglas et al., 1996; Uhler and Emsbo-Mattingly, 2006; Emsbo-Mattingly et al., 2006). For example, the "bulk" ratios of n-C_{18}/n-C_{30} or percent total petroleum hydrocarbons (TPH) in the diesel range (C_{10}–C_{25}/C_{10}–C_{44}) will undoubtedly decrease with evaporative weathering. Alternatively, these same ratios would increase if an HFO had been cut with diesel fuel #2. Other indices may be affected by biological degradation (e.g., n-C_{17}/Pr) or water-washing (e.g., %2-ring PAH/total PAH).

On the other hand, the selected PAH (e.g., C2-dibenzothiphenes/C2-phenanthrenes [D2/P2]) and biomarker indices (e.g., moretane/hopane) are largely considered independent of weathering on an environmental timescale. Therefore, it is proposed that these and other such indices are most suitable for correlating a spilled oil to candidate sources. Among the many potential source ratios, those of greatest value are the ones that clearly differentiate the candidate source oils. Of course, chemical reasonableness in the properties of spilled and candidate source oils should always be considered. In other words, all chemical features (and not only the PAH and biomarker indices) must also support any "positive" correlation based on diagnostic source ratios (Stout et al., 2001).

The impact of oil weathering (e.g., evaporation, solubilization, and biodegradation) on source ratio stability should be considered before it is used in an oil spill investigation (Douglas et al., 1996; Bost et al., 2001; Stout et al., 2001; Daling et al., 2002a; Uhler and Emsbo-Mattingly, 2006). For most waterborne oil spill situations, the spilled oil is primarily influenced by evaporation and volatilization (NRC, 1985); therefore, n-C_{15} and greater hydrocarbons will be less affected. Oil exposed for longer periods of time in the environment will be further altered by biodegradation (NRC, 1985; Prince, 2002). Under severe biodegradation conditions, even the most refractory source ratios can be distorted (Bost et al., 2001); however, these conditions are rarely observed in oil spill events.

Selection of the PAH and biomarker diagnostic indices (e.g., Table 8-3 and Table 8-4) is largely based upon "genetically" significant (i.e., source-specific) variables known to occur among crude oils from different geologic basins. These "genetic" differences are largely carried forward during refining (Peters et al., 1992) and, therefore, can still provide diagnostic information on petroleum products (e.g., a fuel oil). Much of the knowledge of which PAH and biomarker indices are truly "genetically" significant comes from the oil exploration and production geochemical liter-

Table 8-3 List of Common Quantitative Indices Useful in Oil Spill Investigations. This List Is Not Intended to Be All Inclusive

Bulk Indices[1]	*PAH Indices*[1]	*Biomarker Indices*[1]
Isoprenoid Indices	**% Ring Number**	**Sterane Indices**
norpristane/pristane	%[2-ring PAH/Σtotal PAH]	%[$C_{27}\beta\alpha$ diasterane (S/S+R)]
n-C_{17}/pristane	%[3-ring PAH/Σtotal PAH]	%[$C_{29}\alpha\alpha\alpha$ steranes (S/S+R)]
n-C_{18}/phytane	%[4- to 6-ring PAH/Σtotal PAH]	%[$C_{29}\alpha\beta\beta$(R+S)/total C_{29} steranes]
Pristane/Phytane	%[S-PAH/Σtotal PAH]	%[$C_{27}\alpha\beta\beta/C_{27}$–$C_{29}$ $\alpha\beta\beta$]
Boil Range Indices	**Methyl-Phenanthrene Indices**	%[$C_{28}\alpha\beta\beta/C_{27}$–$C_{29}$ $\alpha\beta\beta$]
n-C_{18}/n-C_{30}	MPI 1 = 1.5(2MP + 3MP)/(P + 1MP + 9MP)	%[$C_{29}\alpha\beta\beta/C_{27}$–$C_{29}$ $\alpha\beta\beta$]
%[ΣTPHC$_{10}$–C$_{25}$/ΣTPHC$_{10}$–C$_{44}$]	MPI 2 = 3(2MP)/(P + 1MP + 9MP)	**Triterpane Indices**
CPI (odd $C_{25–33}$/(even $C_{26–34}$) alkanes	MPR = 2MP/1MP	$C_{23}+C_{24}/C_{28}+C_{29}$ tricyclics
	Dibenzothiophene Source Indices	$C_{28}+C_{29}$ tricyclics/Hopane
	MDR = 4MDBT/1MDBT	T_s/Hopane
	DBT/P (Dibenzothiophene/Phenanthrene)	Moretane/Hopane
	D/P	BNH/Hop (28,30-bisnorhopane/Hopane)
	D2/P2	25NH/Hop (25-norhopane/Hopane)
	D3/P3	BNH + 25NH/Hopane
	Naphthobenzothiophene Source Indices	Nor/Hop (Norhopane/Hopane)
	NBT2/C1	Ol/Hop (Oleanane/Hopane)
	NBT3/C2	$C_{29}\alpha\beta/C_{29}T_s$ Hopane
	Retene/Total C4-phenanthrenes	%[C_{31} Hopane (S/S+R)]
	%[Retene/P4]	%[C_{32} Hopane (S/S+R)]
	Chrysene Profile	%[C_{35} Hopanes/ΣC_{30}–C_{35} Hopanes]
	%[C0/Chrysene total]	Hopane/$C_{29}\alpha\alpha\alpha$ 20R Sterane
	%[C1/Chrysene total]	C_{24} tricyclic/C_{26}(S+R) tricyclics
	%[C2/Chrysene total]	Ts/Tm
	%[C3/Chrysene total]	**Triaromatic Sterane (TAS) Source Indices**
	%[C4/Chrysene total]	C_{28},20S TAS/C_{28},20R TAS
	Miscellaneous PAH Source Indices	C_{26},20R+C_{27},20S TAS/C_{27},20R TAS
	FL/PY (BBF + BKF)/BAP	Hop/C_{28},20S TAS
	BBF+BKF/C C/FP1	
	PER/C IP/GHI	
	BBF/BKF PER/BAP	
	BF/FP1 BKF/BAP	
	BBF/BAP BEP/BAP	
	BA/C	

[1] Diagnostic indices may be expressed as A/B or A/(A + B).

ature (Peters and Moldowan, 1993; Peters et al., 2005a, 2005b). These ratios in crude oils are known to be related to (1) the thermal maturity of the source rocks that gave rise to a crude oil, (2) the type of organic matter present in the source rock (e.g., terrestrial versus marine), and/or (3) the degree of weathering that may have occurred in the crude oil reservoir prior to oil production. Understanding the meaning of these ratios, and if they are in any way susceptible to changes due to environmental weathering (evaporation, water-washing, or biodegradation) following an oil spill, certainly needs to be considered (see Chapter 1 herein). However, in this respect, the use of diagnostic ratios based upon biomarker concentrations is particularly useful since these compounds are highly resistant to weathering over environmental timescales (Prince et al., 1994; see also Chapter 11 herein). Finally, although the alkylated PAH compounds are not generally as resistant to weathering as the

Table 8-4 Diagnostic Ratios and Precision-Based Statistics for a Spilled Oil and Two Candidate Source Oils (A and B). All Data are Based Upon the Absolute Concentrations of Individual Compound or Compound Groups Measured by GC/MS as per the Revised Nordtest Method (Daling et al., 2002)

*Diagnostic Ratio**	*Spill Oil (Rep 2)*	*Spill Oil (Rep 1)*	*Spill Oil (Rep 3)*	*Spill Mean*	*Std. Dev.*	*Percent RSD***	*95% CI****	*98% CI****	*Source A*	*Source B*
%27Ts	59	56	57	57	1.5	2.7	3.8	6.1	56	60
%28ab	22	21	18	20	2.1	10.2	5.2	8.4	12	20
%25nor30ab	3	3	2	2	0.5	18.3	1.1	1.8	1	3
%29ab	31	34	31	32	1.7	5.4	4.3	7.0	34	32
%C29Ts	17	18	17	17	0.6	3.3	1.4	2.3	17	18
%30d	9	10	11	10	0.9	8.5	2.1	3.4	13	12
%27dia	46	47	48	47	1.0	2.1	2.5	4.0	48	49
%29aaS	43	49	45	46	3.1	6.7	7.6	12.3	44	44
%29bb	51	53	51	52	1.2	2.2	2.9	4.6	52	49
%27bbSTER	39	38	39	39	0.6	1.5	1.4	2.3	36	40
%28bbSTER	28	30	28	29	1.2	4.0	2.9	4.6	30	27
%29bbSTER	33	34	33	33	0.6	1.7	1.4	2.3	35	33
%TA21	52	50	55	52	2.5	4.8	6.3	10.1	60	54
%TA26	37	36	40	38	2.1	5.5	5.2	8.4	32	39
%TA27	54	56	55	55	1.0	1.8	2.5	4.0	49	56
%D2/P2	31	32	27	30	2.6	8.8	6.6	10.6	17	29
%D3/P3	30	29	25	28	2.6	9.4	6.6	10.6	17	27
%D3/C3	71	77	72	73	3.2	4.4	8.0	12.9	63	75
%2MP/1MP	45	46	47	46	1.0	2.2	2.5	4.0	53	45
%4MD/1MD	76	76	77	76	0.6	0.8	1.4	2.3	85	77

* See Daling et al. (2002) for definition of specific diagnostic ratios.
** Relative percent standard deviation.
*** Confidence Interval as per Student's t test (Harris, 1995).

biomarkers, certain ratios of select PAHs have proven to be reliable and stable for fingerprinting purposes, even in the face of significant evaporative and biodegradative weathering (e.g., C2-dibenzothiophenes/C2-phenenathrenes, Page et al., 1995; Douglas et al., 1996; Uhler and Emsbo-Mattingly, 2006). The chemical compounds used in these ratios must have similar physical and chemical properties such that they both weather at the same rates in the environment and retain the original oil source ratio (Overton et al., 1981).

It is important to realize that the indices listed in Table 8-3 should not be considered all-inclusive or appropriate for all oil spill studies. Sometimes it may be prudent to include a certain index that is recognized as particularly diagnostic of the spilled oil. Acquiring the breadth of concentration data (*n*-alkanes, isoprenoids, PAHs, and biomarkers) described here allows one that flexibility. In addition, the analytical and interpretive methods discussed earlier are not limited only to marine oil spill situations. The basic technical approach and diagnostic indices can also be applied to groundwater/soil investigations where nonaqueous phase liquids (NAPL) and petroleum contaminated soils are present.

8.3.4 Source Identification Protocols for Quantitative Fingerprinting Data

Diagnostic ratios calculated from quantitative or semiquantitative data need to be compared in an unbiased manner. Several methods of increasing sophistication are available for this. First, as we noted earlier, several of the *Exxon Valdez* investigations relied upon quantitative

fingerprinting data to reach conclusions regarding the sources of hydrocarbons in the sediments of Prince William Sound. These studies often relied upon simple but diagnostic *x-y* cross-plots of selected diagnostic ratios to demonstrate similarities and differences among a large number of samples (e.g., Bence et al., 1996). Plots relying upon the C2-dibenzothiophenes/C2-phenanthrenes, C3-dibenzothiophenes/C3-phenanthrenes, oleanane/hopane, 28,30-bisnorhopane/hopane, and Ts/Tm ratios were particularly useful distinguishing cargo oil from other hydrocarbon sources in the Sound. A simple example of this approach to comparing samples is shown in Figure 8-2. Double-ratio cross-plots have the advantage of being able to compare a large number of samples simultaneously, but these have the disadvantage of necessarily relying on only two ratios at a time.

Second, more recently, the Nordtest organization developed a revised oil spill identification protocol based on semiquantitative or quantitative chemical fingerprinting that considers as many diagnostic ratios as is appropriate simultaneously (Daling et al., 2002a). The revised Nordtest protocol relies on the linear correlation of diagnostic ratios calculated from semi- or fully quantitative data generated using GC/FID and GC/MS. In this method, the degree of statistical correlation depends on the analytical precision of the diagnostic ratios, as measured by the 95% and 98% confidence intervals derived from the "Student's t" statistical tool (Harris, 1995) determined from replicate (usually triplicate) analyses. Three levels of correlation were suggested in the revised Nordtest protocol:

Positive match — all diagnostic ratios within the 95% confidence interval.

Probable match — all diagnostic ratios within the 98% confidence interval.

Nonmatch — any key diagnostic ratio outside the 98% confidence interval.

An example of the revised Nordtest protocol involved a mystery oil slick discovered in an offshore oil production area containing numerous oil production platforms, pipelines, and barge traffic. Samples of the slick and numerous candidate sources (production streams from different platforms, pipelines, and barges) were collected for analysis. Since the crude oils produced in the area were generated from oil source rock strata containing similar ancient organic matter that had been heated to a similar extent, there was not a great deal of heterogeneity among the candidate oils. In other words, they were genetically similar, which could have confounded an assessment of the slick's source if quantitative data were not available (as described in Section 8.2.1.3).

The mystery spill oil was prepared and analyzed in triplicate via GC/FID and GC/MS as described in Section 8.3.2. Several candidate source oils were also prepared and analyzed in the same manner. Table 8-4 contains the quantitative results for 20 diagnostic ratios determined for the spill oil (in triplicate) and two of the candidate source oils. The relevant statistics for the spill oil (e.g., mean, standard deviation, relative percent standard deviation, and Student's t confidence interval) were calculated according to Daling et al. (2002a). Important among these is the very good analytical precision evident in the %RSD less than 10 for most diagnostic ratios (Table 8-4).

According to the revised Nordtest method's linear regression protocol (Daling et al., 2002a), the diagnostic ratios for the mean spill oil are plotted against the diagnostic ratios for the two candidate source oils on simple *x-y* plots (Figure 8-3). If there was a perfect match between the spill oil and any candidate source oil, each point (representing a particular diagnostic ratio from Table 8-4) would fall perfectly along a straight line. Such perfect matches are rarely observed, inevitably due to some degree of analytical precision error. Two examples among the various candidate source oils analyzed were selected for demonstrating the revised Nordtest's elegantly simple regression approach — Source A and Source B (Figure 8-3). Figure 8-3A clearly shows that Source A has several diagnostic ratios that plot

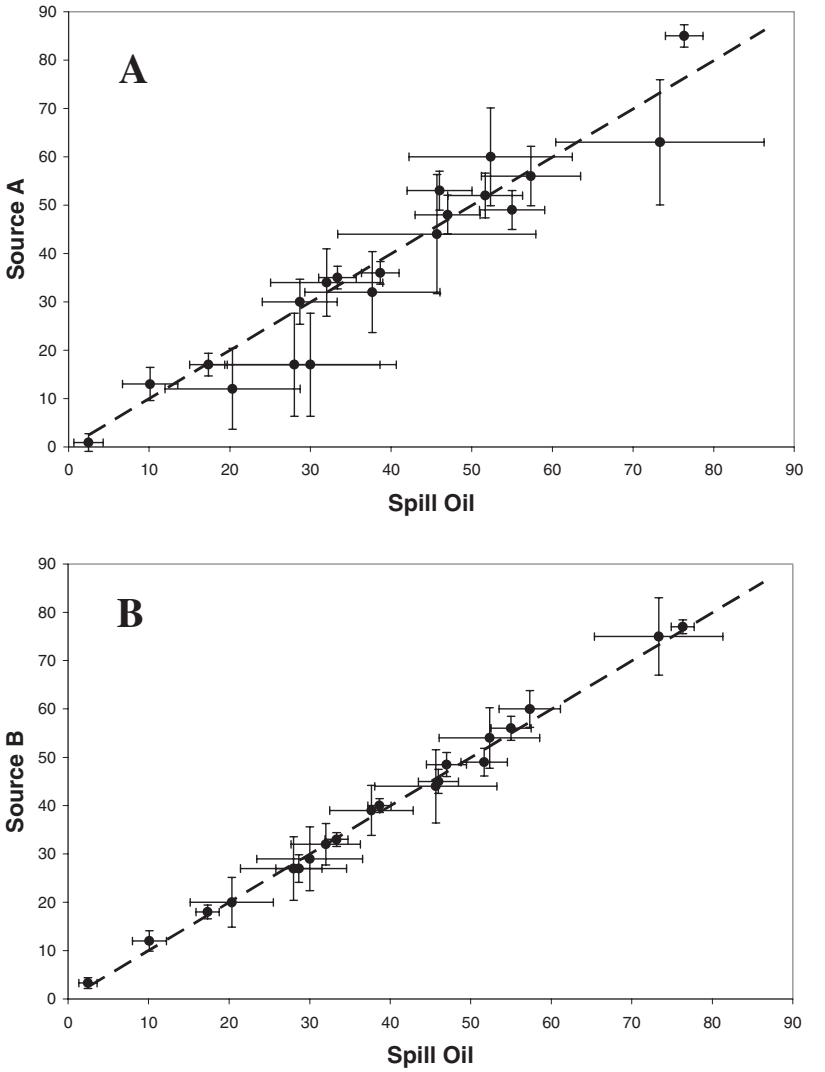

Figure 8-3 Examples of linear regression analysis as per the revised Nordtest oil spill identification method for the data shown in Table 8-4. (A) Example of "non-match" between spill oil and Source A based upon 98% confidence interval. (B) Example of "match" between spill oil and Source B based upon 98% confidence interval.

beyond the 98% CI indicating that, according to the revised Nordtest method's recommended criteria, Source A is a "nonmatch" to the spilled oil. On the other hand, Figure 8-3B shows that Source B has all 20 of the diagnostic ratios within the 95% CI of the spilled oil's ratios indicating that, according to the recommended criteria, Source B is a "positive match" to the spill oil.

Despite the overall genetic similarity among the candidate source oils, the Source B oil was determined to be the only "positive match" among the available source candidates. This provided a strong chemical basis upon which to conclude it was the likely source of the mystery slick. Importantly, there was no opportunity of bias entering into this statistically based interpretation of the quantitative data.

Although the revised Nordtest protocol offers an entirely statistical and objective correlation approach, one potential issue with this

approach is that laboratories with exceptional analytical precision for their replicates will have a smaller error and therefore more rigorous match criteria than laboratories with a lower analytical precision. Therefore, it is possible that the less precise laboratory will identify a "positive match" when the more precise lab will not. Another potential issue of the revised Nordtest approach is that because it relies on correlation between only two oils at a time (typically a spill versus a potential source), it can have difficulty dealing with large datasets and with spill samples influenced by mixtures (Stout et al., 2005). The revised Nordtest protocol served as the starting point for the emerging CEN protocol described in Chapter 7.

One issue related to the revised Nordtest protocol is that its simple graphical technique (e.g., Figure 8-3) can become cumbersome if a large number of candidate oils need to be compared to a spilled oil. In such a situation, we develop a correlation matrix based on some standard similarity index, such as the correlation coefficient (r or r^2), based on a statistical comparison of the diagnostic ratios among the oils. A simple ranking of these indices for each candidate source oil versus the spilled oil — from highest to lowest — helps identify those candidate source oils most highly correlated with the spill. The revised Nordtest graphical approach to matching can then focus on these most highly correlated oils in order to determine if they meet appropriate match criteria.

Some additional comments on the use of similarity indices in oil spill correlation studies are warranted. While similarity indices, such as r or r^2, can be used effectively to rank the relative similarity of field samples, it is important to recognize the limitations of correlation analysis for source identification purposes. First, it is important to use quantitative data (e.g., analyte concentrations) with known levels of precision and accuracy. It is particularly helpful to include replicate results and confidence intervals whenever possible. Second, as described, the diagnostic parameters should resist environmental weathering. The indiscriminant use of all analyte data is commonly confounded by nonlinear changes in analyte concentrations due to the preferential weathering of more volatile or biodegradable constituents. Dimensionless concentration ratios of recalcitrant analytes (e.g., biomarkers and 3- to 6-ring PAHs) generally serve well as source indicators in correlation analysis. Third, diagnostic ratios should be equally weighted so that a small number of source indicators with a wide dynamic range do not obscure the statistical importance of indicators with a narrow dynamic range. Specifically, the analyte ratios should be Z-score or range-normalized before running the correlation analysis. Fourth, the selection of levels of significance for correlation coefficients should not be established based on an absolute scale (e.g., "$r^2 > 0.75$ is a source signature match"). Among other factors, the level of significance for the correlation coefficient should take into account the degree of similarity (e.g., genetic similarity) among multiple source area samples and degrees of difference among samples collected from nonimpacted reference areas. Finally, the results of any correlation analysis should consider the potential for mixing. Oil from the same source may accumulate different features as it accumulates hydrocarbons from the ambient environment or from independent releases. Accordingly, the application and results of any simple correlation analysis based on some similarity index must be evaluated for chemical reasonableness and checked on an individual basis using qualitative and quantitative fingerprinting techniques.

An alternative means of analyzing a large suite of quantitative fingerprinting data simultaneously is to employ some multivariate statistical method, such as principal component analysis (PCA) or partial least-squares analysis (e.g., Aboulkassim and Simoneit, 1995; Burns et al., 1997; Stout et al., 2001; Lavine et al., 2001; Mudge, 2002; Christensen, 2002; Christensen et al., 2004, 2005; Li et al., 2004). These methods and their advantages are more fully described in Chapter 9. The primary benefit of these methods is that — as long as the input data consist of data whose precision

is known and that is considered source-diagnostic (and not prone to weathering effects) — multivariate analysis of quantitative data can produce objective and confident positive and negative correlations. Obviously, such analyses cannot be performed on qualitative data.

As discussed earlier, correlations based on simple double-ratio cross plots, revised Nordtest-like correlation methods, or even multivariate analyses methods cannot be performed in a vacuum. In other words, diagnostic ratios evaluated by any of these methods need to consider all of the available data before a "positive match" is achieved. Also, keep in mind that a chemical match does not unequivocally mean that the correlated candidate source oil is the only possible source oil. Each spill's circumstances, including any other site information such as location of the spill, water current direction, operational practices, or chemical information such as the degree of weathering and trace biomarker differences, may provide the additional evidence required to isolate the spill source.

8.4 Unraveling Mixed Source Oils Using Quantitative Fingerprinting Data

8.4.1 Two-Component Mixing Models

The one-to-one linear regression protocol described in the revised Nordtest method and exemplified in Figure 8-3, or multivariate analyses, can work extremely well if there is no likelihood of multiple sources mixing to produce a mystery spill. However, there are oil spills in which there is a significant prespill oil component (background) or mixtures of different oils spilled simultaneously that need to be recognized and accommodated. The revised Nordtest protocol (Daling et al., 2002a) acknowledges that "mixed source signals" could be difficult to recognize and that when mixtures are suspected, "a simple linear mixing algorithm could be developed to explain the oil spill." Simple linear mixing of a two-component mixture is straightforward and has been used in oil field studies (e.g., Kaufman et al., 1990; Peters and Fowler, 1992). However, since the diagnostic ratios relied on in the revised Nordtest method do not necessarily respond linearly to mixing (see discussion in Section 8.3.1), the development of any linear mixing algorithm *requires* the use of the absolute concentration of the compound(s) in the algorithm. In other words, the mixing of the diagnostic ratios themselves does not consider that the individual compound concentrations will likely vary between the two mixed oils. For example, Table 8-5 contains a set of data in which the absolute concentrations of two biomarkers (A and B) and two PAHs (C and D) are measured in two oils and the ratios A/(A + B) and C/(C + D) calculated in each oil. Theoretical mixtures of these two oils will yield a set of diagnostic ratios with values intermediate to starting ratios. When the ratios are calculated based on the absolute concentrations of the biomarkers and PAHs, the resulting ratios accurately reflect the ratios in their mixtures. However, suppose the original ratios of the candidate source oil were based upon relative abundances (e.g., peak heights). Mixing the original ratios results in set of diagnostic ratios that do not accurately reflect the mixtures due to the varying concentrations of these biomarkers and PAHs in the oils. Thus, in an oil spill investigation where mixtures of different oils (or oils with background contamination) are under investigation, fully quantitative data are necessary to develop accurate mixing models. The need for concentration data is heightened when the analysis involves data generated by multiple laboratories or data generated over a long period of time. This is a very strong argument for the need to obtain fully quantitative (absolute concentration) data for individual chemicals in oil spill fingerprinting studies.

In other situations, it is possible that the candidate sources collected for a mystery oil spill investigation do not fully represent the population of candidate sources. For example, consider the situation where a vessel's slop tank is empty at the time (after the spill event) samples are being collected, but it was not

Table 8-5 Example Data for Selected Biomarkers for Two Oils Demonstrating How the Absolute Analyte Concentrations are Needed to Calculate Accurate Diagnostic Ratios in Theoretical Mixtures. Ratios Calculated from the Original Ratios, as Might Be Available from Peak Areas (but not Concentrations), Do Not Mix Linearly Due to Concentration Differences in the Two Oils. Therefore, Absolute Concentrations Are Necessary to Develop Accurate Mixing Models

	OIL 2 (mg/kg)	*Theoretical Mixtures of OIL 1 and OIL 2*								*OIL 1 (mg/kg)*
(OIL1 : OIL2)	**(0 : 100)**	**(10 : 90)**	**(20 : 80)**	**(30 : 70)**	**(40 : 60)**	**(50 : 50)**	**(60 : 40)**	**(70 : 30)**	**(80 : 20)**	**(90 : 10)**	**(100 : 0)**
Biomarker A	63.9	59.1	54.4	49.7	45.0	40.3	35.5	30.8	26.1	21.4	16.6
Biomarker B	83.3	78.2	73.1	68.1	63.0	58.0	52.9	47.9	42.8	37.8	32.7
PAH C	80.9	74.5	68.2	61.9	55.6	49.2	42.9	36.6	30.3	23.9	17.6
PAH D	403.8	399.4	394.9	390.5	386.0	381.6	377.1	372.7	368.2	363.8	359.3
Diagnostic Ratios in Mixtures Based Upon Absolute Analyte Concentrations											
%A/(A + B)	43.4	43.1	42.7	42.2	41.6	41.0	40.2	39.2	37.9	36.1	33.7
%C/(C + D)	16.7	15.7	14.7	13.7	12.6	11.4	10.2	8.9	7.6	6.2	4.7
Diagnostic Ratios in Mixtures Based Upon Simple Mixtures of Original Source Ratios											
%A/(A + B)	43.4	42.4	41.5	40.5	39.5	38.6	37.6	36.6	35.7	34.7	33.7
%C/(C + D)	16.7	15.5	14.3	13.1	11.9	10.7	9.5	8.3	7.1	5.9	4.7

when the spill occurred. Since the vessel's slop tank might have contained heavy fuel oil (HFO), lubricating oil, or other oils from the vessel, it is possible that some mixture of these oils was present in the slop tank at the time of the spill. Mathematically reconstructing a theoretical suite of possible "slop tank" mixtures — based on the individual oil that might have been present in the slop tank — requires fully quantitative (absolute concentration) data for target PAH and biomarkers in each of the individual end-member candidate oils. We demonstrate this approach in two case studies that involved a mixed source signal.

8.4.2 Case Study 1

In this case study, an oil spill was discovered adjacent to a docked marine crude oil tanker — which may not seem like a mystery. However, quantitative GC/MS analysis of PAH and biomarkers in three candidate source oils collected from the vessel's cargo tanks, HFO fuel tanks, and bilge tanks showed them each to be "nonmatches" to the spilled oil according to the revised Nordtest criteria (Figure 8-4). One of the three candidate source oils — the crude oil cargo — was very different from the spill and could be eliminated for further consideration. However, of the 10 diagnostic ratios utilized (A through J; Figure 8-4), the spilled oil exhibited multiple diagnostic ratios that appeared intermediate between the HFO fuel and bilge oil. For example, inspection of Figure 8-4A shows that the bilge oil exhibited several diagnostic ratios that were higher than observed in the spill oil (e.g., ratios B and F) and another that was lower than in the spill oil (ratio C). Inspection of Figure 8-4B shows that the HFO exhibits diagnostic ratios that were lower than observed in the spill oil (e.g., ratios B and F). Thus, at this point in the investigation, there was a suspicion that the spilled oil might have been derived from a mixture of the bilge and HFO oils, but this needed to be proven.

In order to evaluate the candidacy of a mixed oil release, the PAH and biomarker absolute concentration data were evaluated using a linear mixing algorithm in which the concentrations of each PAH and biomarker analyte were calculated in a series of theoretical, two-component mixtures according to the following equation:

$$(\text{Conc. A} \times \%\text{bilge}) + (\text{Conc. A} \times \%\text{HFO}) \quad \text{(Eq. 8-1)}$$

Recalculation of the 10 diagnostic ratios (A through J) based on the concentrations of each ratio's numerator and denominator chemicals

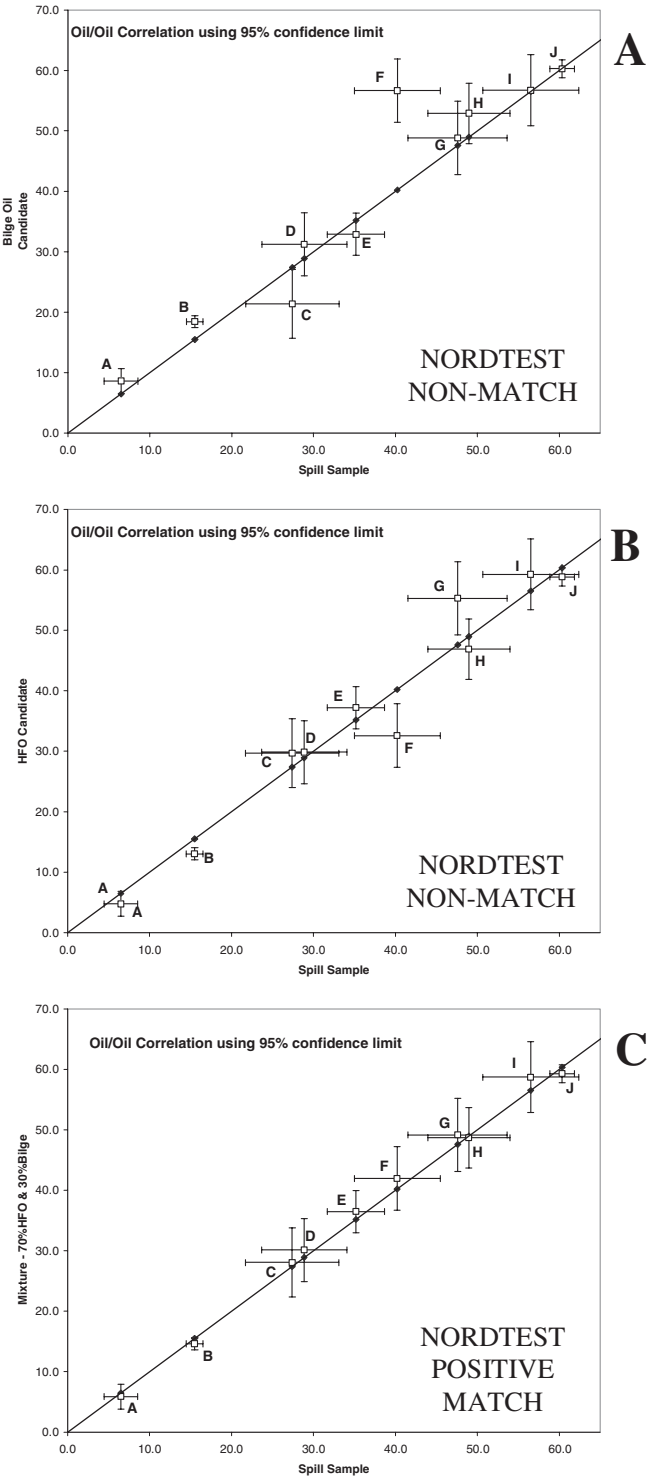

Figure 8-4 Revised Nordtest linear regression plots for ten diagnostic ratios (A-J) in the mystery spill oil and two candidate source oils from case study 1. (A) comparison of bilge oil with spilled oil resulting in a non-match, (B) comparison of HFO oil with spilled oil resulting in a non-match, (C) comparison of theoretical mixture of bilge and HFO with spilled oil resulting in a positive match.

using this two-component mixing algorithm showed that a mixture of 70%HFO and 30%bilge oil provided a "positive match" to the spilled oil according to the revised Nordtest's linear regression protocol, the result of which is shown in Figure 8-4C. Thus, it was reasonably concluded that the suspect vessel had discharged an approximately 70:30 mixture of HFO and bilge oil and was thereby likely responsible for the mystery spill. This conclusion would have been entirely speculative if the absolute concentration data were not available and the two-component mixing model could not be developed.

8.4.3 Case Study 2

Another recent case study evaluated mixed oil signals using quantitative chemical fingerprinting data combined with the revised Nordtest method. In this case, a commercial vessel carrying large quantities of intermediate and heavy fuel oils (IFO and HFO, respectively) and a small amount of marine diesel fuel (MDO) was accidentally grounded. The grounding resulted in IFO and HFO being released to the marine environment. During the subsequent investigation, samples of the IFO and HFO from the vessel as well as field samples potentially containing oil(s) released from the grounded vessel were collected. The objective of the study was to determine if any field samples (mousse and tarballs) collected after the release were "positive matches" to the source oils from the grounded vessel. Although this objective may seem like a simple task, standard interpretive approaches to source identification require careful considerations when mixtures of oils potentially are present (e.g., bilges; Section 8.2.1.4).

Figure 8-5 shows the GC/FID chromatograms, PAH concentration histograms, and tricyclic and pentacyclic triterpane biomarker fingerprints for the IFO and HFO collected from the vessel's fuel tanks after the grounding. The disparate nature of these two residual fuel oils is reflected in their distinct GC/FID chromatograms (Figure 8-5A-B), which can be distinguished well based on a qualitative inspection of the hydrocarbon patterns. The GC/FID chromatographic differences are best expressed by the wider range of *n*-alkanes and bimodal unresolved complex mixture (UCM, e.g., higher molecular weight) present in the HFO relative to the IFO.

Quantitative chemical fingerprinting analysis revealed a particularly significant difference in the PAH composition of the two fuels, namely the substantially lower C2-dibenzothiophenes/C2-phenanthrenes and C3-dibenzothiophenes/C3-phenanthrenes (D2/P2 and D3/P3, respectively) ratios in the IFO relative to the HFO (Figure 8-5C-D). The environmental stability and source character of these ratios are well established (Douglas et al., 1996, Page et al., 1995), indicating that these fuels contain different levels of sulfur-bearing PAH. In addition, quantitative biomarker analysis showed markedly lower biomarker concentrations of triterpanes in the IFO relative to the HFO, the absence of C28 and C29 tricyclic terpanes in the HFO, and markedly different "triplet" distributions (i.e., C26 tricyclic and C24 tetracyclic terpanes; see inset) and 30-norhopane/hopane signatures (Figure 8-5E-F).

Because of the obvious differences in the GC/FID chromatograms, PAH distributions, and triterpane concentrations/distributions, qualitative fingerprinting would likely be sufficient to distinguish these two source oils from one another, presumably even after moderate weathering has occurred. Thus, it is clear that the two fuels carried by the grounded vessel were chemically distinct from one another.

Figure 8-6 depicts the GC/FID, PAH, and triterpane biomarker fingerprints for two apparently impacted field samples — a mousse sample and a tarball sample — collected after the grounding. The on-site responders confidently linked the mousse sample to the release based on visual observations after the spill. Prior to obtaining the vessel fuel oil samples, the mousse sample was initially used as a representative source sample to identify other mousse and tarball samples in the area. Direct qualitative comparison of the GC/FID

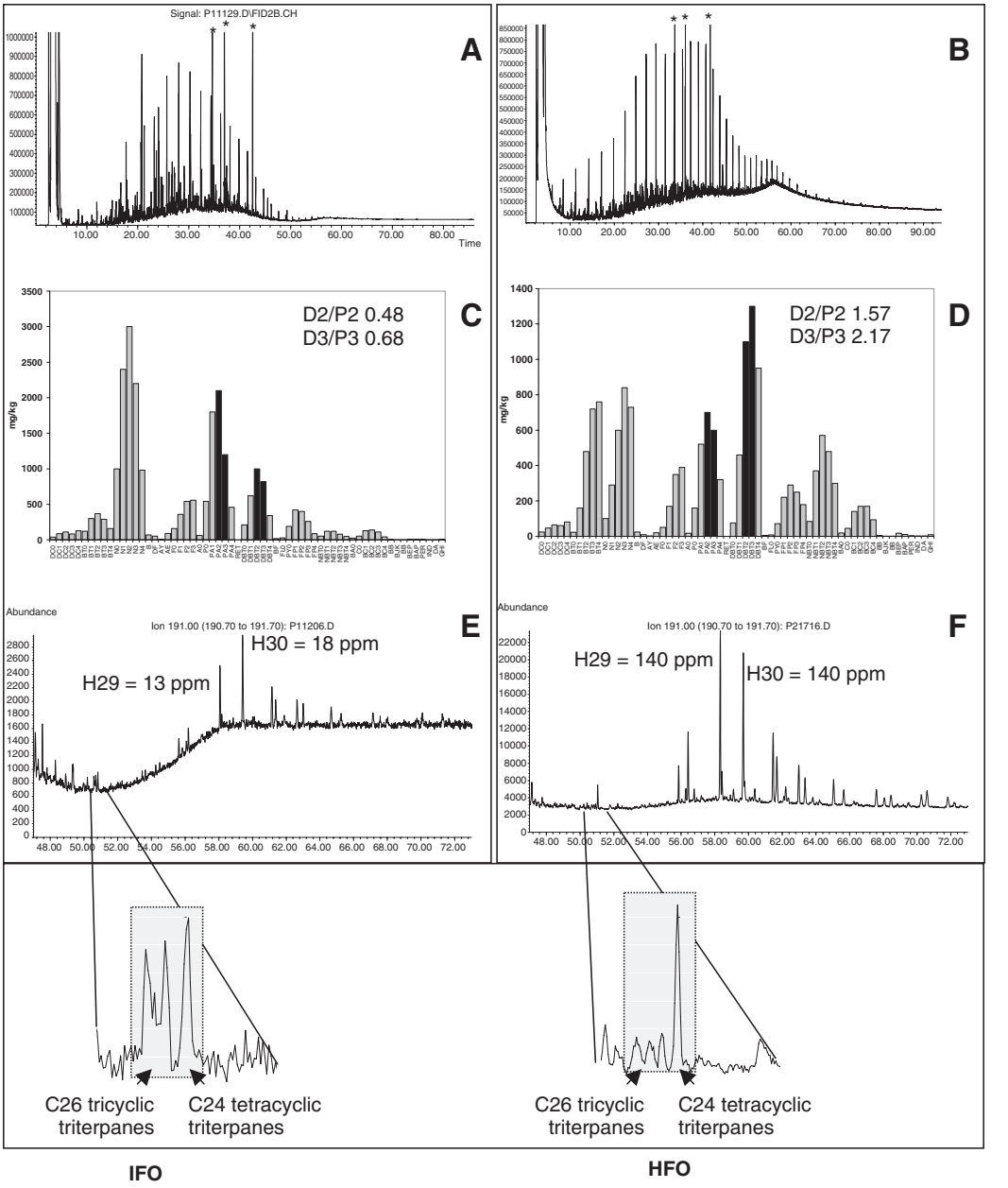

Figure 8-5 Comparison of IFO and HFO source fuels oil fingerprints from a marine oil spill from case study 2. (A) GC/FID chromatogram for IFO, (B) GC/FID chromatogram for HFO, (C) PAH concentration histogram for IFO, (D) PAH concentration histogram for HFO, (E) Partial *m/z* 191 triterpane extracted ion profile for IFO, and (F) Partial *m/z* 191 triterpane extracted ion profile for HFO. H29 = Norhopane, H30 = Hopane * = internal standards.

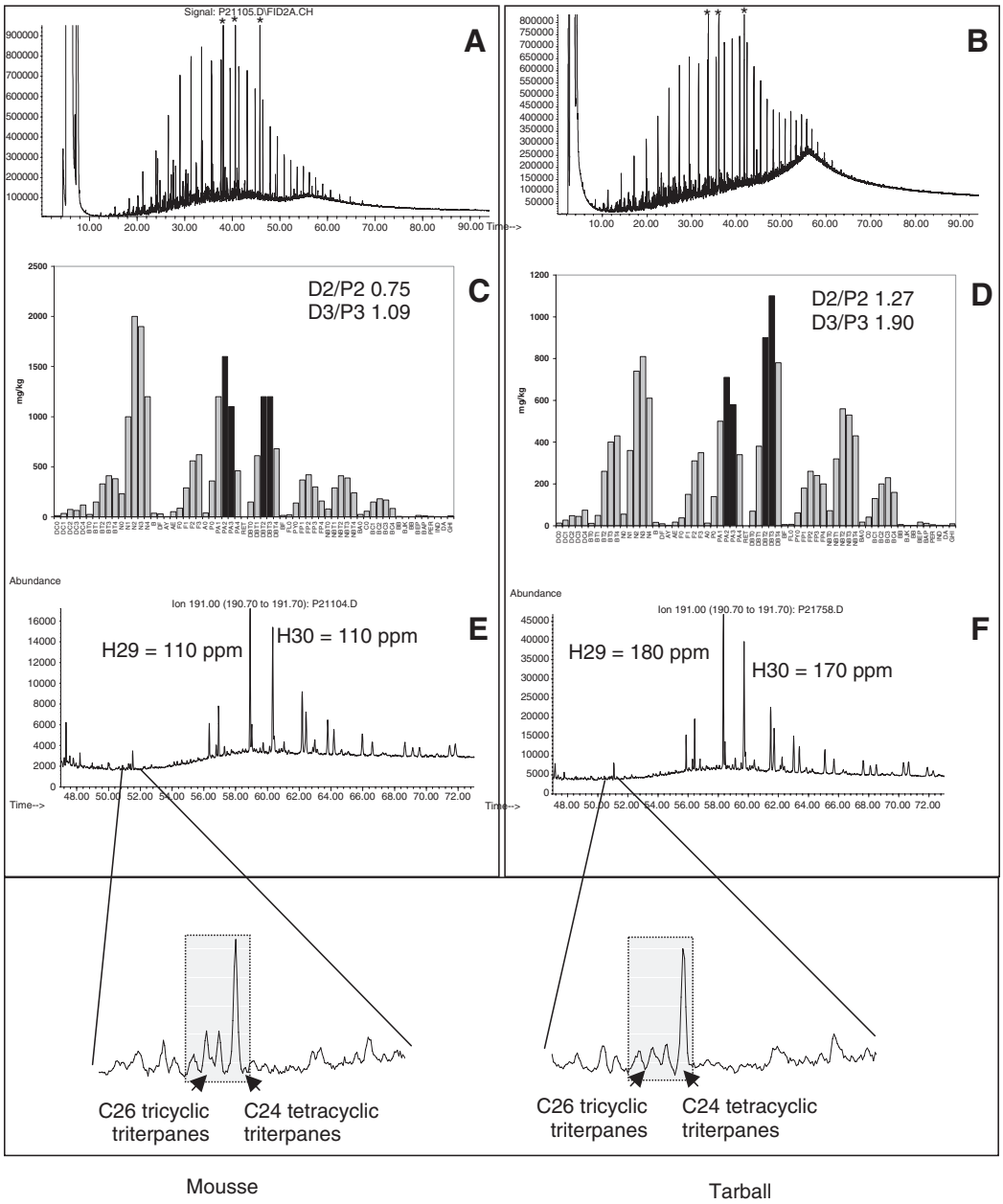

Figure 8-6 Comparison of two field samples (mousse and tarball) associated with case study 2 and the vessel containing the fuels shown in Figure 8-5. (A) GC/FID chromatogram for the mousse sample (B) GC/FID chromatogram for the tarball sample, (C) PAH concentration histogram for the mousse sample, (D) PAH concentration histogram for the tarball sample, (E) Partial *m/z* 191 triterpane extracted ion profile for the mousse sample, and (F) Partial *m/z* 191 triterpane extracted ion profile for the tarball sample. H29 = Norhopane, H30 = Hopane * = internal standards.

chromatograms revealed that the two field samples (and others not discussed in this case study) exhibited heterogeneities in the distribution of *n*-alkanes and in the shape of the UCM hump (Figure 8-6A-B). It would be difficult to say with certainty that these differences were not attributable to variable degrees of weathering of a single oil. For example, evaporative losses of compounds from the HFO (Figure 8-5B) might yield a GC/FID fingerprint comparable to the tarball sample (Figure 8-6B). However, the quantitative PAH data indicated that the mousse and tarball samples were not derived from a single oil. For example, the PAH distributions and D2/P2 and D3/P3 ratios clearly showed that the field samples ($n > 100$) contained varying abundances of sulfur-containing PAHs (Figure 8-6C-D). This result raised some questions as to the actual source(s) of these and other field samples, particularly for samples collected in areas not previously identified as oil-impacted. The questions posed by the on-site responders were (1) are there multiple sources of oil contamination at this site? (2) why are the field samples chemically different from the two source fuel oils? and (3) what tools are available to *confidently* identify spill- and nonspill-related field samples?

Direct qualitative comparison of the two source oils from the grounded vessel (Figure 8-5) and the two field samples (compare Figures 8-5 and 8-6) indicates that neither of the sources would likely match either of the mousse or tarball samples. (This was later confirmed quantitatively using the revised Nordtest method; see ahead.) Although there were general chemical similarities between the two source oils and the field samples (e.g., all were seemingly comprised of a broad-boiling, *n*-alkane-enriched oil), there were significant chemical differences (e.g., D2/P2 and D3/P3 ratios) that could *not* be explained by simple weathering. Because the grounding and oil spill occurred in an area of heavy commercial shipping traffic, where many of the vessels may have carried generally similar fuels, further investigation as to the other potential source(s) of oil or pre-existing oil in the field samples was warranted.

A suite of the source-specific, diagnostic ratios was developed for these oils (Table 8-6) and was investigated with the objective of determining if either of the vessel's fuel oils had contributed to the oil in the field samples. These ratios were generated from quantitative analysis of the source oils and field samples as described in Section 8.3.2 and were selected

Table 8-6 Diagnostic Ratios for the IFO and HFO Source Oils and Field Samples

Source Ratio	*Abbr.*	*Source Samples*		*Field Samples*		*Calculated Mixtures*		*Field Sample (Unrelated)*	
		IFO	*HFO*	*Mousse Sample*	*Tarball Sample*	*Mousse Sample Mixture (50:50)*	*Tarball Sample Mixture (12:88)*	*Tarball #2 Sample*	*Tarball #2 (100:–2)*
100 × [Pr/(Pr+Ph)]	a	63.89	59.41	58.56	63.06	62.91	60.90	61.96	63.91
100 × [odd C25–C33/(odd C25–C33+even C26–C34)]	b	61.29	57.87	53.31	56.77	58.48	57.97	62.52	61.57
100 × [DBT/(DBT+P0)]	c	28.00	31.91	29.41	33.01	28.93	30.71	15.49	27.98
100 × [D2/(D2+P2)]	d	32.26	61.11	42.86	55.90	42.86	55.52	31.71	31.98
100 × [D3/(D3+P3)]	e	40.59	68.42	52.17	65.48	54.09	64.82	30.16	40.15
100 × [Ts/(Ts+Hop)]	f	16.67	14.63	17.29	16.67	14.87	14.67	15.44	16.96
100 × [More/(More+Hop)]	g	0.00	7.89	6.86	8.11	7.06	7.77	16.36	–1.29
100 × [NHop/(NHop+Hop)]	h	41.94	50.00	50.00	51.43	49.20	49.88	43.90	40.51
100 × [C31 S/(C31 S+C31 R)]	i	56.50	55.38	57.84	56.90	55.52	55.41	67.01	56.65
100 × [C32 S/(C32 S+C32 R)]	j	62.04	58.67	56.25	58.10	59.19	58.75	49.62	62.38
100 × [C35 S+R/((C35 S+R)+ΣC30–C35 Hopanes)]	k	10.89	13.48	10.95	12.10	13.14	13.43	0.00	10.57

based on their source-specific properties, their reported resistance to environmental weathering, and their low relative percent difference (RSD < 10% indicates a high degree of analytical precision) for the triplicate analyses of the mousse. Thus, these diagnostic ratios would meet the general criteria for diagnostic ratios (e.g., Stout et al., 2001). The diagnostic ratios (DR) were calculated using the general formula

$$DR = 100 \times [A/(A + B)] \qquad \text{(Eq. 8-2)}$$

The diagnostic ratio results range from 0 to 100, per the revised Nordtest (Daling et al., 2002). The diagnostic ratios for the two source oils, mousse, and tarball samples are provided in Table 8-6 — along with some calculated diagnostic ratios for theoretical mixtures of the IFO and HFO sources and a nonspill-related tarball sample (Tarball #2) — that is discussed later in this section.

The diagnostic ratios for each field sample were plotted versus the same ratios for each of the vessel's source oils (Figures 8-7 and 8-8) relative to the parity line, per the revised Nordtest method. The parity line represents the 1 : 1 source/sample ratio, where identical oils would plot close to or directly on this line. The 98% confidence limits derived from the triplicate analysis of the mousse sample are plotted with each data point (*x* and *y* ranges) to reflect the associated analytical variability. Samples where all ratios fall within this confidence level are identified as a probable match using the recommended revised Nordtest criteria discussed in Section 8.3.4.

Analysis of these figures indicates that the sulfur-related diagnostic ratios (e.g., ratio $d = [D2/D2 + P2] \times 100$, ratio $e = [D3/D3 + P3] \times 100$) are clearly different from the IFO and HFO for both the mousse and the tarball samples, respectively (i.e., they plot off the parity line). This was observed earlier in Figures 8-5 and 8-6, where D2/P2 and D3/P3 ratios in the field samples were markedly different than those observed in the field samples. Given that the mousse sample was visually identified by the responder as being spill-related, and the fact that many of the shoreline tarball samples also failed the revised Nordtest criteria, we evaluated the possibility that the field samples were indeed various mixtures of the two source oils.

Several lines of physical and chemical evidence indicated that many of the field samples were likely mixtures of the two source oils. Specifically,

1. Visual observations by on-site responders identified areas of shoreline near the grounded vessel that were clearly related to the spill event (e.g., mousse sample).
2. Qualitative and quantitative chemical analysis of the source oils and field samples suggested that although most of the field samples were a chemical "nonmatch," they appeared to have some genetic biomarker similarities (e.g., C26 tricyclic and C24 tetracyclic terpanes).
3. The range of source ratios observed in the field samples was mostly bounded by the IFO (low) and HFO (high).
4. The primary differences in the revised Nordtest correlation plots appeared to be driven by the sulfur ratios (e.g., D2/D2 + P2), in which the field samples' ratio was intermediate to the two source oils.

Although similarity was readily apparent between the candidate source oils and the spill samples, the revised Nordtest approach was not able to fully unravel the true source of a mixed product that formed the spill sample (e.g., Figure 8-7). Similarly, this approach could not be used to correlate the individual spilled products to tarballs (formed both during the spill and from pre-existing spills) discovered on the nearby coastline. To resolve this issue, a two-component mixing model composed of various proportions of the two source oils was developed using published methods (Page et al., 1995; Burns et al., 1997; Arouri and McKirdy, 2004; Kauffman et al., 1990). Because the sulfur ratios (e.g., D2/D2 + P2) in the source oils and field samples accounted for a majority of the variance in the samples (Figures 8-7 and 8-8), these ratios were used to develop a simple mathematical model to estimate the percentage of each

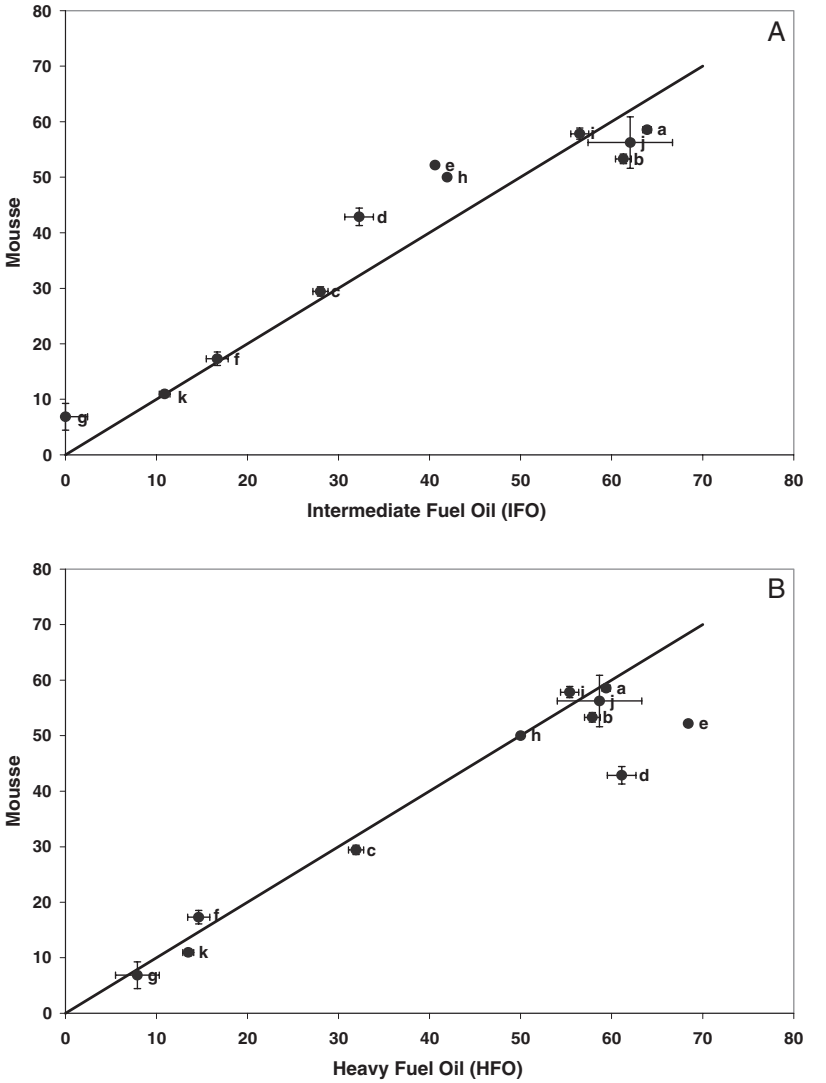

Figure 8-7 Revised Nordtest linear regression plots for eleven diagnostic ratios (a-k) in the mousse sample and two candidate source oils (IFO and HFO) from case study 2. (A) Comparison of IFO with the mousse sample resulting in a non-match, (B) comparison of HFO oil with the mousse sample resulting in a non-match.

source oil within the field sample(s). A new set of diagnostic ratios was then calculated based on the percentage of each source oil in the field sample and used to evaluate the sensitivity of the revised Nordtest method to resolve the identity of the field samples.

8.4.3.1 Mixing Model Case Study 2

Commercial merchant vessels often store three types of fuel; IFO, HFO, and MDO (EPA, 1999). IFO and HFO represent the majority of the vessel's fuel cargo (typically hundreds of thousands of gallons), whereas smaller amounts of MDO are commonly used only to operate auxiliary generators on the vessel. Uhler et al. (2006) (Chapter 10 in this book) document the significant variability in the ratios of D2/P2 and D3/P3 in HFOs. Again, in this case, a significant difference in these ratios between the cargo fuels was recognized (as discussed ealier) and used to develop the spill

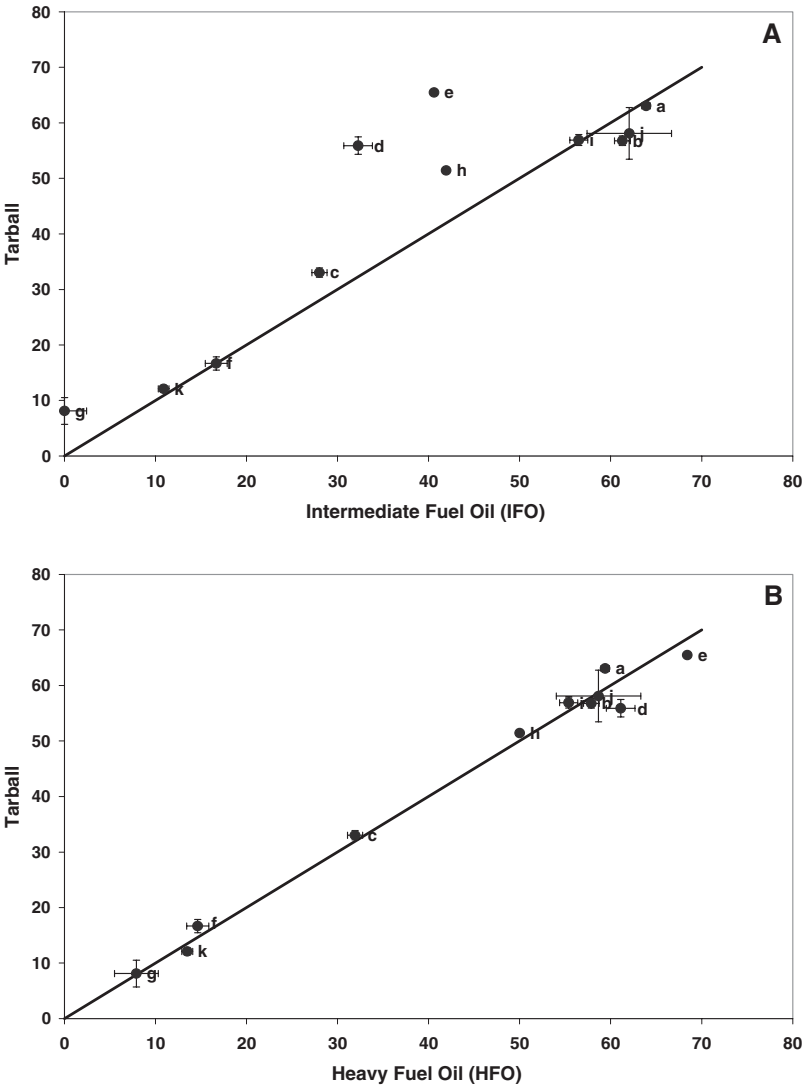

Figure 8-8 Revised Nordtest linear regression plots for eleven diagnostic ratios (a-k) in the tarball sample and two candidate source oils (IFO and HFO) from case study 2. (A) Comparison of IFO with the tarball sample resulting in a non-match, (B) comparison of HFO oil with the tarball sample resulting in a non-match.

mixing model. Rather than simply use the C2-dibenzothiophene concentration in the source oils, source ratios of C2-dibenzothiophenes/C2-phenanthrenes (D2/P2) were used because

- These source ratios vary widely in oils depending on its crude source (Page et al., 1995).
- These source ratios are stable over a wide range of weathering (Douglas et al., 1996).
- These source ratios are less variable than concentrations because they are internally normalized (to C2-phenanthrene) in the field samples.

The mixing model ratios are calculated from *absolute concentrations* of C2-dibenzothiophenes and C2-phenanthrenes in the source oils. Some error may be introduced because the MDO was not included in the model; however, given the relatively small volume of MDO on the vessel and the lack of a clear MDO signal in the field samples, any bias would be minor. In addition, there was some

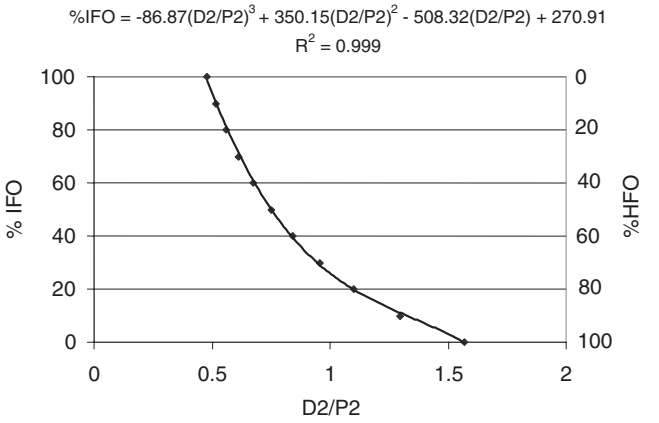

Figure 8-9 Plot of the D2/P2 versus percent IFO and HFO from the grounded vessel. This equation was used with the D2/P2 ratios measured in the field samples to calculate the % IFO and % HFO in the field samples. These percentages were then used with the field sample quantitative data to calculate mixed oil source ratios that were used in the revised Nordtest method (e.g., Fig. 8-10).

information from the on-site responders that the MDO tank had been pumped out. Figure 8-9 is a plot of the D2/P2 ratios versus %IFO in a mixture of the two source oils (IFO and HFO). The equation is expressed by the following third-order polynomial equation:

$$\%\text{IFO} = -86.87(\text{D2/P2})^3 + 350.15(\text{D2/P2})^2 - 508.32(\text{D2/P2}) + 270.91 \quad \text{(Eq. 8-3)}$$
$$r^2 = 0.9995$$

The %HFO is simply 100% – %IFO. Once the IFO : HFO proportion for each field sample is determined, a new suite of diagnostic ratios (Table 8-6) is calculated from the absolute concentration data using an equation comparable to Eq. (8-1) (but for the two source oils in Case Study 2).

The best fit for the mixing model in Figure 8-9 is not linear due to the differences in concentrations of the two compound groupings within the source oils (C2-dibenzothiophenes, C3-dibenzothiophenes, Page et al., 1995). This fact re-emphasizes why absolute concentrations are necessary in developing mixing models. For example, the slope of the mixing model curve is flatter with higher percentages of HFO because it takes very little HFO to have a large influence on the D2/P2 ratio in the mixture. Conversely, the slope of the curve is greater with higher IFO concentrations because a given change in D2/P2 requires substantially more IFO relative to the HFO. For the mousse sample, the best calculated relative percentage of source oils based on Eq. (8-3) and the D2/P2 of the sample was a 50 : 50% mixture of IFO and HFO. For the tarball sample, the calculated relative percentage of source oils was a mixture of 12 : 80% IFO and HFO. These mixtures were consistent with other chemical factors observed in the sample including UCM distribution, alkane and PAH distribution, tricyclic and tetracyclic triterpanes distribution, and biomarker distributions.

Figure 8-10A and B is a plot of the calculated diagnostic ratios based on the mixing model versus the quantitative mousse (Figure 8-10A) and tarball sample data (Figure 8-10B). The recalculation of the mixed source(s) based on the mixing model significantly improved the fit for both the mousse and tarball samples and strongly indicated that the field samples were indeed variable mixtures of the source oils. For the mousse sample, the improvement was dramatic (Figure 8-7A) and brought the sample very close to the revised Nordtest method's recommended positive match criteria (Figure 8-10A). The primary drivers (as discussed earlier) were the sulfur ratios as exhibited by source ratios *c* ([DBT/DBT + P0] × 100), *d* ([D2/D2 + P2] ×

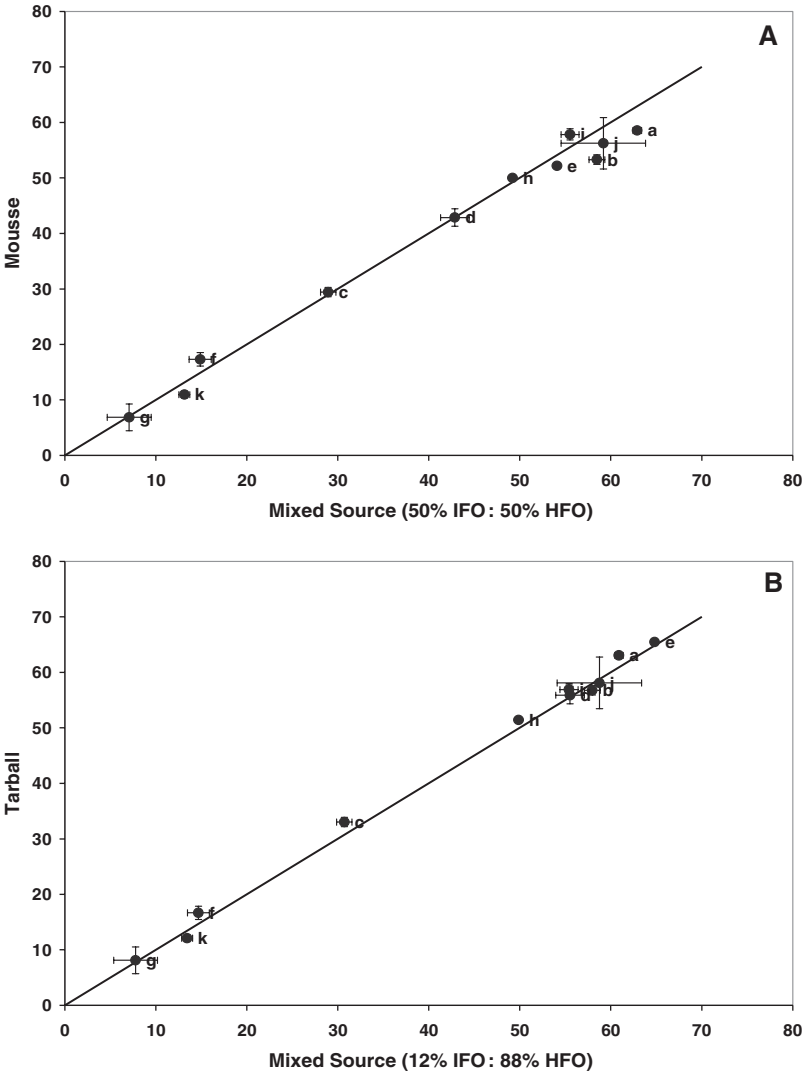

Figure 8-10 Revised Nordtest linear regression plots for eleven diagnostic ratios (a-k) in the mousse (A) and tarball (B) samples versus the associated mixed source oils. The parity correlation was improved for both samples when the quantitatively determined mixed source ratios were used. Based on this analysis and considering the additional variability introduced by the mixing model but not reflected in the 98% confidence interval, a source match with the mixed oils is likely.

100), and *e* ([D3/D3 + P3] × 100), which were shifted closer to the parity line for a mixture of 50% IFO and 50% HFO. A similar shift closer to the parity line was observed in the tarball sample for a mixture of 12% IFO and 88% HFO. These results are also consistent with the GC/FID, PAH, and biomarker interpretive results. This model was then applied to a suite of oil, mousse, and tarball field samples collected during the ongoing field investigation months after the spill to confidently resolve spill-related oiling from nonspill-related (preexisting) contamination.

The revised Nordtest method (and the emerging CEN method; Chapter 7) provides a series of specific guidelines to reliably identify

mystery oil sources within single component oil spills. However, when mixtures of oils are present, the simple application of this approach may produce a substantial number of false negative results. Under these circumstances, as demonstrated within this case study, the application and source resolution of the revised Nordtest method are greatly enhanced through the development of source oil mixing models. Observation of a significant improvement in parity line correlation between the mixed source and the field sample is a primary indicator of mixing model efficacy (Figure 8-10). Unlike the revised Nordtest approach, however, the error associated with the correlation analysis is not only related to the analytical variability (precision) determined from replicates of a single oil, it also includes the added potential variability introduced by the mixing model itself. This potential variability is a function of how representative the source oils are to the bulk fuel oil, and if all possible sources have been collected from the vessel. This additional error has not been incorporated into the current study although it is expected to increase the 98% confidence intervals associated with the calculated mixtures (e.g., *x*-axis). Principal component and least-squares methods (Burns et al., 1997) based on all of the source ratios (not simply D2/P2) may substantially reduce the variance associated with the simple two-component mixing model described earlier. These multivariate methods can also more readily accommodate mixtures involving more than two potential sources. In addition to the calculation of mathematical mixes, source oil mixtures could be physically prepared in the laboratory and triplicate analyses performed to more accurately estimate this potential increase in variance. For the purposes of this study, the post-model improvements (Figure 8-10) relative to the sample/mixed-source parity line is strong evidence that the mixed oils are representative of the spilled oils and should be considered, at a minimum, as probable matches based on the revised Nordtest method. In addition to field observations and the qualitative inspection of the hydrocarbon patterns discussed previously, the samples exhibiting a high degree of correlation with the parity line constitute a defensible match with the source oils in this case.

The mixing model approach used in this case study provides an example of a useful method for exploring a range of source signatures derived from binary mixtures of potential source oils. Evaluating a broader range of potential source signatures provides a higher level of confidence for the match/nonmatch conclusion and minimizes the false negatives encountered with simpler correlations of unicomponent sources to spill samples composed of mixtures of spilled oils. The examination of an infinite number of mixtures bounded by the sources dramatically increases the accuracy of the fingerprinting method when a nonspill-related sample is identified. Within oil spill response situations, it is the identification of nonspill-related oils that are most problematic to the environmental regulators; therefore, the methods used to reach such a conclusion must be highly defensible and supported by the chemical data.

Figure 8-11 is a mixed-source ratio plot for another field sample from this case study (Tarball #2, Table 8-6), but it was not related to the known oil spill under investigation. The initial revised Nordtest plots relative to the IFO and HFO indicated a clear nonmatch, which was not unexpected given that we had identified the possibility of source oil mixtures at the site. The analysis was then expanded to examine all possible combinations of the source oils, including the best combination identified with the mixing model described above. In all cases there was no combination of the two source oils that could explain the source ratio chemistry observed in the sample. The results presented in Figure 8-11 for Tarball #2 show the disparity between the tarball and the best possible source oil mixture. These results demonstrate that the sample is unrelated to the oil spill and *not* a false negative. In addition to the revised Nordtest results, the conclusion was confirmed by direct examination of the GC/FID, PAH, and biomarker distributions and abundances.

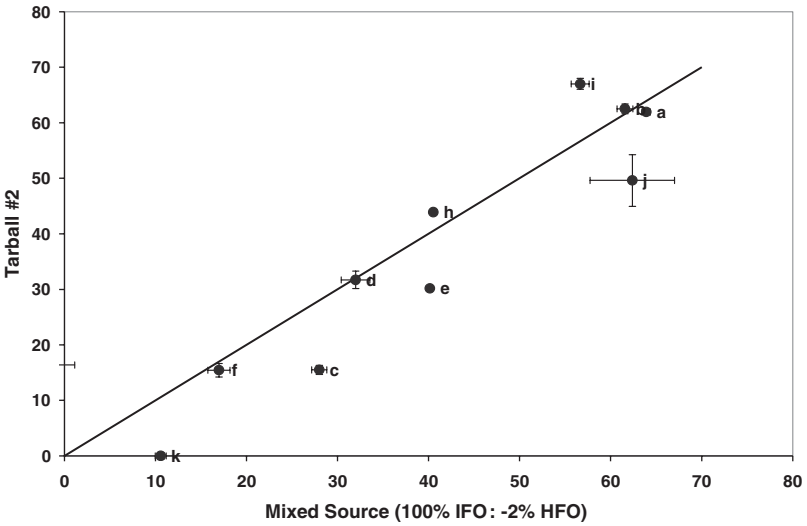

Figure 8-11 Revised Nordtest linear regression plots for eleven diagnostic ratios (a-k) in the tarball #2 sample versus the associated mixed source oil. No improvement in parity correlation was observed over the entire range of possible mixtures including the mixed source presented in this figure that was calculated based on the D2/P2 ratio in the sample. This data combined with GC/FID, PAH and biomarker signatures is conclusive evidence that this sample is not related to the oil spill discussed in case #2.

Based on our experience and as discussed earlier, chemical differences between potential oil spill sources at a site can be relatively small or vastly different. The parity plot analysis discussed earlier may be used to (1) evaluate the potential for source mixtures, (2) rank the degree of correlation between field samples relative to source oils, and (3) define a subset of source oils most likely related to the release. This refined subset of possible sources can then be evaluated in more detail based on the preponderance of the evidence such as the spill location, site history, process chemistry, degree of weathering, and subtle but unique differences in the GC/MS and GC/FID chemical signatures.

8.5 Summary

The chemical fingerprinting approaches available for the identification of oil spill sources and potentially impacted samples fall into two categories: (1) qualitative and (2) quantitative. The qualitative approach relies upon visual comparison of various chromatographic fingerprints and is exemplified by the ASTM D3328 and D5739 methods. The quantitative fingerprinting methodology relies on measurements of the concentrations (relative or absolute) of targeted and diagnostic chemicals, typically PAH and biomarkers, and a subsequent statistical or other numerical analysis of various diagnostic parameters calculated from these concentrations.

The quantitative approach is represented by the revised Nordtest methodology (Daling et al., 2000a), discussed in this chapter, and the emerging CEN methodology (Chapter 7 herein). The quantitiative approach is preferable for most oil spill investigations since the means of interpretation are more objective and robust in the sense that they facilitate numerical comparison of diagnostic details and reduce interpretation bias. The quantitative approaches have a much better opportunity to account for and recognize the effects of weathering, and the subtle but real differences among genetically or otherwise qualitatively similar oils than qualitative fingerprinting methods. However, quantitative approaches such as the revised Nordtest or emerging CEN methods still can be confounded when mixing

affects the chemical fingerprints of samples under consideration. Mixing is an important consideration in any oil spill investigation involving multiple source oils or when a spilled oil mixes with any pre-existing hydrocarbons in the ambient environment, both of which will alter the chemical fingerprint of a spilled oil.

To account for the effects of mixing, absolute concentration data are needed for the samples under consideration. The effects of mixing can be mathematically accommodated with numerical mixing models so long as fully quantitative data (i.e., absolute concentration data) are available. When absolute concentration data are available for end-members, a numerical mixing model can predict the character of intermediate mixtures. The use of the revised Nordtest (or emerging CEN) sample-to-sample correlation approach in conjunction with numerical mixing models can improve the chemical comparison between spilled oil(s), candidate source oils, and potentially impacted environmental samples. These numerical mixing models must be based on absolute (concentration) quantitative chemical data since diagnostic ratios calculated from semiquantitative data do not mix linearly due to absolute concentration differences between samples. As with any forensic study, a weight of evidence approach should be used and traditional analysis of all the available chromatographic and analytical data should be carefully examined to ensure that any conclusions regarding correlation between samples is consistent with the mixing model and subsequent statistical analysis.

References

Aboulkassim, T.A.T. and B.R.T. Simoneit, Petroleum hydrocarbon fingerprinting and sediment transport assessed by molecular biomarker and multivariate statistical-analyses in the Eastern harbor of Alexandria, Egypt. *Marine Pollution Bulletin*, 1995, **30**(1), 63–73.

Arouri, K.R. and D.M. McKirdy, The behavior of aromatic hydrocarbons in artificial mixtures in Permian and Jurassic end-member oils: Application to in-reservoir mixing in the Eromanga Basin, Australia. *Organic Geochemistry*, **36**, 105–113.

ASTM, Oil spill identification by gas chromatography and positive ion electron impact low resolution mass spectrometry. ASTM D-5739–00. American Society for Testing and Materials International, W. Conshohocken, PA, 2000.

ASTM, Standard Test Methods for Comparison of Waterborne Petroleum Oils by Gas Chromatography. ASTM D-3328–90. American Society for Testing and Materials International, W. Conshohocken, PA, 1990.

Bence, A.E. and W.A. Burns, Fingerprinting hydrocarbons in the biological resources of the *Exxon Valdez* spill area. In P.G. Wells, J.N. Butler, and J.S. Hughes (eds.), Exxon Valdez *Oil Spill: Fate and Effects in Alaskan Waters*. ASTM STP 1219. American Society for Testing and Materials, W. Conshohocken, PA, 1995, pp. 84–140.

Bence, A.E., K.A. Kvenvolden, and M.C. Kennicutt, Organic geochemistry applied to environmental assessments of Prince William Sound, Alaska, after the *Exxon Valdez* oil spill — a review. *Organic Geochemistry*, 1996, **24**(1), 7–42.

Boehm, P.D., G.S. Douglas, W.A. Burns, P.J. Mankiewicz, D.S. Page, and A.E. Bence, Application of petroleum hydrocarbon chemical fingerprinting and allocation techniques after the *Exxon Valdez* oil spill. *Mar. Pollut. Bull*., 1997, **34**(8), 599–613.

Bost, F.D., R. Frontera-Suau, T.J. McDonald, K.E. Peters, and P.J. Morris, Aerobic biodegradation of hopanes and norhopanes in Venezuelan crude oil. *Organic Geochemistry*, 2001, **32**, 105–114.

Burns, W.A., P.J. Mankiewicz, A.E. Bence, D.S. Page, and K.R. Parker, A principal component and least squares method for allocating polycyclic aromatic hydrocarbons in sediment to multiple sources. *Environ. Toxicol. Chem*., 1997, **16**(6), 1119–1131.

Christensen, J.H., Application of multivariate data analysis for assessing the early fate of petrogenic compounds in the marine environment following the *Baltic Carrier* oil spill. *Polycyclic Aromatic Compounds*, 2002, **22**(3–4), 703–714.

Christensen, J.H., A.B. Hansen, G. Tomasi, J. Mortensen, and O. Andersen, Integrated methodology for forensic oil spill identification. *Environ. Sci. Technol.*, 2004, **38**(10), 2912–2918.

Christensen, J.H., G. Tomasi, and A.B. Hansen, Chemical fingerprinting of petroleum biomarkers using time warping and PCA. *Environmental Science & Technology*, 2005, **39**(1), 255–260.

Daling, P.S., L.-G. Faksness, A.B. Hansen, and S.A. Stout, Improved and standardized methodology for oil spill fingerprinting. *Environ. Forensics*, 2002a, **3**, 263–278.

Daling, P.S. and L.-G. Faksness, Laboratory and reporting instructions for the CEN/BT/TF 120 Oil Spill Identification Round Robin test — May 2001. SINTEF Report STF66 A02027, 2002b.

Douglas, G.S., A.E. Bence, R.C. Prince, S.J. McMillen, and E.L. Butler, Environmental stability of selected petroleum hydrocarbon source and weathering ratios. *Environ. Sci. Technol.*, 1996, **30**(7), 2332–2339.

Douglas, G.S., R.C. Prince, E.L. Butler, and W.G. Steinhauer, The use of internal chemical indicators in petroleum and refined products to evaluate the extent of biodegradation. In R.E. Hinchee, B.C. Alleman, R.E. Hoeppel, and R.N. Miller (eds.), *Hydrocarbon Bioremediation.* Ann Arbor, MI, Lewis Publishers, 1994.

Douglas, G.S. and A.D. Uhler, Optimizing EPA methods for petroleum contaminated site assessments. *Env. Test. Anal.*, 1993, **2**, 46–53.

Emsbo-Mattingly, S.D., A.D. Uhler, S.A. Stout, G.S. Douglas, K.J. McCarthy, and A. Coleman, Determining the source of PAHs in sediment. *Land Contamination & R*, 2006, **14**(2).

EPA, In-use marine diesel fuel. Engine programs and Compliance Division, Office of Mobile Sources, U.S. Environmental Protection Agency. EPA420-R-99-027, 1999.

Faksness, L.-G., P.S. Daling, and A.B. Hansen, Round-robin study — oil spill identification. *Environ. Forensics*, 2002, **3**(3/4), 279–292.

Federal Register, National Oil and Hazardous Substances Pollution Contingency Plan — Final Rule 40, CFR Parts 9 and 300, 1994.

General Oceanics, Product and Services Catalog. 1295 N.W. 163rd Street, Miami, FL, 33169, USA, 2006.

Harris, D.C., *Quantitative Chemical Analysis*, 4th ed. New York, NY: W.H. Freemantle & Company, 1995.

Henry, C.B., P.O. Roberts, and E.B. Overton, Advancing forensic chemistry of spilled oil: Self-normalizing fingerprint indexes. *In Proc. 1997 Intl. Oil Spill Conf.* 1997, Fort Lauderdale, FL, March 1997, pp. 936–937.

Hwang, R.J., D.K. Baskin, and S. Teerman, Allocation of commingled pipeline oils to field production. *Org. Geochem.*, 2000, **31**, 1463–1474.

ISO, Standard 8217, *Specifications of Marine Fuels*, in Petroleum products — Fuels (class F). International Organization for Standardization (ISO), Geneva, Switzerland, 1996.

Kauffman, R.L., A.S. Ahmed, and R.J. Elsinger, Gas chromatography as a development and production tool for fingerprinting oils from individual reservoirs: Applications in the Gulf of Mexico. *In Proc. 9th Annual Research Confe. Society of Econ. Paleontologists and Mineraologists* (D. Schumaker and B.F. Perkins, eds.), New Orleans, 1990, pp. 263–282.

Kvenvolden, K.A., F.D. Hostettler, P.R. Carlson, J.B. Rapp, C.N. Threlkeld, and A. Warden, Ubiquitous tar balls with a California-source signature on the shorelines of Prince William Sound, Alaska. *Environ. Sci. Technol*., 1995, **29**, 2684–2694.

Lehr, W., C. Barker, and D. Simecek-Beatty, New developments in the use of uncertainty in oil spill forecasts. *In Proc. 22nd Artic and Marine Oilspill Program* (AMOP), Calgary, Canada, June 1999, pp. 271–284.

Lavine, B.K., D. Brzozowski, A.J. Moores, C.E. Davidson, and H.T. Mayfield, Genetic algorithm for fuel spill identification. *Analytica Chimica Acta.*, 2001, **43**, 233–246.

Marine Oil Spill Response Corporation (MSRC), Hydrocarbon chemistry analytical methods for oil spill assessments. Marine Spill Response Corporation, Research and Development Program. MSRC Technical Report Series. 95-032, 1995, pp. 77–114.

McCaffrey, M.A., H.A. Lagarre, and S.J. Johnson, Using biomarkers to improve heavy oil reservoir management: An example from the Cymric Field, Kern County. *Am. Assoc. Petrol. Geol. Bull.*, 1996, **80**(6), 904–919.

Mudge, S.M., Reassessment of the hydrocarbons in Prince William Sound and the Gulf of Alaska: Identifying the source using partial least squares. *Environ. Sci. Technol.*, 2002, **36**, 2354–2360.

NRC (National Research Council), *Oil in the Sea: Inputs, Fates, and Effects*. Washington DC: National Academic Press, 1985.

NRC (National Research Council), *Oil in the Sea III.* Washington DC: National Academic Press, 2003.

Nicolle, G.C., C. Boibien, H.L. ten Haven, E. Tegelaar, and P. Chavagnac, Geochemistry: A powerful tool for reservoir monitoring. Soc. Petrol. Engineers, Special Paper No. 37804, 1997, pp. 395–401.

Nordtest, Nordtest Method for Oil Spill Identification. NT Chem 001. Edition 2. NORDTEST. Esbo, Finland, 1991.

Overton, E.B., J.A. McFall, S.W. Mascarella, C.F. Steele, S.A. Antoine, L.R. Politzer, and J.L. laseter, Identification of petroleum residue sources after a fire and oil spill. *In Proce. 1981 Intl. Oil Spill Conf.* American Petroleum Institute, Washington, DC, 1981, pp. 541–546.

Page, D.S., P.D. Boehm, G.S. Douglas, and A.E. Bence, Identification of hydrocarbon sources in the benthic sediments of Prince William Sound and the Gulf of Alaska following the *Exxon Valdez* oil spill. In P.G. Wells, J.N. Butler, and J.S. Hughes (eds.), Exxon Valdez *Oil Spill: Fate and Effects in Alaskan Waters*. ASTM STP 1219. Philadelphia, PA: ASTM, 1995, pp. 41–83.

Page, D.S., P.D. Boehm, G.S. Douglas, A.E. Bence, W.A. Burns, and P.J. Mankiewicz, Pyrogenic polycyclic aromatic hydrocarbons in sediments record past human activity: A case study in Prince William Sound, Alaska. *Mar. Pollut. Bull.*, 1999, **38**, 247–260.

PERF, *Guidelines for the Scientific Study of Oil Spill Effects* — Draft Preface. Petroleum Environmental Research Forum, 1995.

Peters, K.E. and M.G. Fowler, Applications of petroleum geochemistry to exploration and reservoir management. *Organic Geochemistry*, 1992, **33**, 5–36.

Peters, K.E., G.L. Scheuerman, C.Y. Lee, J.M. Moldowan, R.N. Reynolds, and M.M. Pena, Effects of refinery processes on biological markers. *Energy Fuels*, 1992, **6**(5), 560–577.

Peters, K.E. and J.M. Moldowan, *The Biomarker Guide*. Englewood Cliffs, NJ: Prentice Hall, 1993.

Peters, K.E., C.C. Walters, and J.M. Moldowan, *The Biomarker Guide, Volume 2. Biomarker and Isotopes in Petroleum Exploration and Earth History.* Cambridge, UK: Cambridge University Press, 2005a.

Peters, K.E., C.C. Walters, and J.M. Moldowan, *The Biomarker Guide, Volume 1. Biomarker and Isotopes in the Environment and Human History.* Cambridge, UK: Cambridge University Press, 2005b.

Philp, R.P., *Fossil Fuel Biomarkers: Applications and Spectra*. Methods in Geochemistry and Geophysics, Vol. 23. Elsevier: Amsterdam, The Netherlands, 1985.

Prince, R.C., Biodegradation of petroleum and other hydrocarbons. In G. Bitton (ed.), *Encyclopedia of Environmental Microbiology*. New York: John Wiley, 2002, pp. 2402–2416.

Prince, R.C., D.L. Elmendorf, J.R. Lute, C.S. Hsu, C.E. Haith, J.D. Senius, G.J. Dechert, G.S. Douglas, and E.L. Butler, 17a (H) 21 β (H)-hopane as a conserved internal marker for estimating the biodegradation of crude oil. *Environ. Sci. Technol.*, 1994, **28**(1), 142–145.

Radke, M., D.H. Welte, and H. Willsch, Maturity parameters based on aromatic hydrocarbons: Influence of the organic matter type. *Org. Geochem.*, 1986, **10**, 51–63.

Stout, S.A., G.S. Douglas, A.D. Uhler, K.J. McCarthy, and S.D. Emsbo-Mattingly, Identifying the source of mystery waterborne oil spills — a case for quantitative chemical fingerprinting. *Environmental Claims Jour.*, 2005, **17**(1), 71–88.

Stout, S.A., W.P. Naples, A.D. Uhler, K.J. McCarthy, L.G. Roberts, and R.M. Uhler, Use of quantitative biomarker analysis in the differentiation and characterization of spilled oil. Society of Petroleum Engineers, Inc., Richardson, TX. SPE Publ. No. 61460. Stavanger, Norway, June 2000.

Stout, S.A., A.D. Uhler, and K.J. McCarthy, A strategy and methodology for defensibly correlating spilled oil to source candidates. *Environ. Forensics*, 2001, **2**, 87–98.

Uhler, A.D. and S.D. Emsbo-Mattingly, Environmental stability of PAH source indices in pyrogenic tars. *Bull. Environ. Contam. and Tox*, in press, 2006.

Tissot, B.P. and D.H. Welte, *Petroleum Formation and Occurrence*, 2nd ed. Berlin: Springer-Verlag, 1984.

Urdal, K., N.B. Vogt, S. Pl. Sporstol, R.G. Lichtenthaler, H. Mostad, K. Kolset, S. Nordenson, and K. Esbensen, Classification of weathered crude oils using multimethod chemical analysis, statistical methods and SIMCA pattern recognition. *Mar. Pollut. Bull.*, 1986, **17**(8), 366–373.

Wang, Z. and M. Fingas, Identification of the source(s) of unknown spilled oils. *In 1999 Intl. Oil Spill Conf.* Seattle, WA, Paper #162, 1999.

Wang, Z., M. Fingas, and D.S. Page, Oil spill identification. *J. Chromatogr. A.*, 1999, **842**, 369–411.

Wang, Z., M. Fingas, and L. Sigouin, Characterization and identification of a "mystery" oil spill from Quebec (1999). *J. Chromatogr. A.*, 2001, **909**, 155–169.