Application of PAH and biomarker diagnostic ratios in forensic oil spill identification by the revised Nordtest methodology


[^1]: Department of Environmental Chemistry & Microbiology, National Environmental Research Institute (NERI), Roskilde, Denmark.
[^2]: Maritime Environment Section, Admiral Danish Navy, Århus, Denmark.

Abstract

For some years Denmark, like several other countries, has used forensic oil spill identification based on GC-MS fingerprinting according to the Nordtest method for oil spill identification. With this methodology the chemical fingerprints are usually compared visually rendering the final conclusion subjective and dependent on the analyst’s experience. As a result of a recent project initiated by Nordtest, a revised and improved methodology has incorporated PAH and biomarker diagnostic ratios and statistical evaluation to give more robust and technically defensible conclusions. In Denmark, the National Environmental Research Institute (NERI) is responsible for the forensic oil spill analysis as part of the national marine oil spill preparedness. In this paper we describe the implementation of the revised Nordtest methodology and the application of diagnostic ratios together with statistical evaluation for the forensic identification of a recent marine oil spill in Danish territorial water.

1 Introduction

In Denmark, the National Environmental Research Institute (NERI) has been responsible for analysing and identifying oil spills as part of the national marine oil spill preparedness for more than 15 years. Chemical analyses are considered as an integrated and necessary part of the investigation and prosecution of maritime vessels suspected of violating the Marpol 73/78 convention. Until recently, NERI applied the Nordtest method for oil spill identification [1] as the standard procedure when performing forensic oil spill analyses. In 2000, Nordtest ini-
tiated a project to revise and improve the existing methodology. Besides the Nordic forensic oil spill laboratories, Battelle Laboratories, Duxberry, USA, also participated. This project resulted in a revised methodology as presented in these proceedings [2]. NERI therefore recently adopted this new methodology when carrying out forensic oil spill identification, and in this paper we describe an example on how this new approach has been used in a recent marine oil spill case.

1.1 Revised Nordtest methodology for oil spill identification

With the existing Nordtest method, identity is concluded when no intrinsic differences are identified between spill and suspected source samples; otherwise samples are considered non-identical. The methodology is based on a three-step procedure: GC-FID screening of all available samples from the spill and suspected sources, GC-MS fingerprinting of spill samples and candidate source samples, and a final weathering check, e.g. based on n-alkane distributions as evaluated from the GC-FID screening. Identity/non-identity is basically concluded on a visual comparison of GC-MS fingerprints of spill and suspected source samples, and the weathering check may help evaluate whether any observed differences are inherent or simply caused by weathering or analytical variations. This approach is largely dependent on the analyst’s experience and his subjective evaluation thereby rendering the conclusion weakly technically defensible.

The revised Nordtest method has been improved by introducing a more robust and technically defensible methodology based on objective and statistical evaluation. It still operates as a three-step procedure but includes a series of diagnostic ratios derived from PAH and biomarker GC-MS fingerprints that are correlated statistically within e.g. the 95% confidence limits. Based on the statistical evaluation conclusions are furthermore subjective to a four-level evaluation: positive match, probable match, inconclusive and non-match.

1.2 Case story of marine oil spill

In August 2001, a marine oil spill was observed within the Danish territorial zone, and a helicopter from the Admiral Danish Navy was directed towards the spill to observe, report and collect samples. A foreign cargo tanker heading for the Baltic Sea, was observed in the vicinity of the spill. The next day NERI received two samples from the spill and still a week later another seven oil samples from the suspected tanker collected at a port state control by the authorities of Port Kaliningrad, Russia. NERI was asked to perform a forensic oil spill identification of the spill samples with respect to the suspected source samples.

2 Chemical analysis

According to the revised Nordtest methodology for oil spill identification all samples were prepared for GC-FID and GC-MS analysis [3].
2.1 GC-FID screening

The GC-FID analyses were performed using standard GC equipment and in a way that ensured good quality chromatograms with sufficient resolution of the two n-alkane/isoprenoid pairs, nC17/pristine (Pr) and nC18/phytane (Ph), and of individual n-alkanes up to about nC40. The GC-FID screening of spill and suspected source samples revealed that only one sample from the cargo tank of the suspected tanker could be a candidate source. The GC-FID chromatograms of those two samples are shown in Figure 1. The chromatograms display irregular (i.e. bimodal) but rather similar alkane/isoprenoid distributions together with other characteristic features like the unresolved hump. Based on GC-FID characteristics, both samples seem to originate from heavy fuel oil residues, and thus illegal tank washing could have caused the spill.

![Figure 1: GC-FID chromatograms of a spill sample from the oil slick at sea (A) and a sample from the cargo tank of the suspected tanker (B).](image)

2.2 GC-MS fingerprinting

The second step of the Nordtest methodology is the GC-MS fingerprinting. The analyses were performed using a high-resolution GC-MS instrument running in single-ion recording (SIR) mode. A total of 35 ions including alkanes, homologue series of alkylated PAHs and selected petroleum biomarkers were recorded in two runs for each sample. The spill sample was analysed in triplicates to provide the standard deviations of derived diagnostic ratios. In the following, Figures 2 to 4 show examples of some of the GC-MS ion fragmentograms (fingerprints).

By visual comparison of the three sets of GC-MS ion fragmentograms, only marginal differences are observed for the triterpane/hopane (Figure 2) and 14β(H)-sterane (Figure 3) fingerprints. Most of the other ion fragmentograms recorded for the two samples confirm these observations, and the marginal differences observed might very well be interpreted as caused by differences in actual concentrations, analytical variations or by physical weathering.
Figure 2: GC-MS fragmentogram (m/z 191,18) of spill sample (A) and candidate source sample (B) showing the hopanes (triterpanes) biomarker pattern.

Figure 3: GC-MS fragmentogram (m/z 218,20) of spill sample (A) and candidate source sample (B) showing the 14β(H)-steranes biomarker pattern.

Figure 4: GC-MS fragmentogram (m/z 192,09) of spill sample (A) and candidate source sample (B) showing the methylphenanthrene isomer pattern.

The methylphenanthrene isomer patterns shown in Figure 4, however, display obvious differences. In the source sample (B), the first doublet (3- and 2-methylphenanthrenes) is higher than second doublet (9-/4- and 1-methylphenanthrenes), while the situation is reversed for the spill sample (A). The ratio of the first and second doublet observed for the source sample is typical for heavy fuel oils and probably due to thermodynamic equilibration of the isomers, while the ratio observed for the spill sample is that typical for crude oils [4, 5]. As the 3- and 2-
methylphenanthrenes are probably more sensitive to weathering/degradation [6, 7] than the other isomers, it is difficult to conclude whether the observed difference is real or caused by degradation. The small but distinct methylanthracene peak (MA) observed in both samples also indicates that both samples are sourced from a heavy fuel oil [5].

2.3 Weathering checks

In the original Nordtest method, a weathering check was achieved by correlating relative GC-FID peak heights of the n-alkanes in the two samples. In the revised Nordtest methodology, weathering checks are obtained from correlating normalised peak heights (or areas) of both n-alkanes (from GC-FID) and a suite of parent and alkylated PAHs (from GC-MS). In both cases peaks are normalised relative to a non-weathered compound, e.g. nC25 for n-alkanes and the C30αβ-hopane for PAHs. Examples of the weathering checks are shown by the bar charts in Figure 5.

![Bar charts showing weathering checks](image)

Figure 5: Weathering checks performed by correlating n-alkanes (A, norm. peak heights) and PAH homologue groups (B, norm. peak areas).

The removal of the more evaporative and soluble compounds is expected as part of natural physical weathering processes [6]. However, the n-alkane weathering check shows the opposite trend as the relative peak heights of lower-boiling compounds are higher in the spill sample than in the source sample. The PAH weathering check displayed similar differences as some PAHs are higher in the spill sample than in the source, e.g. the fluorene, dibenzothiophene and chrysene homologue series. Hence, these differences cannot be attributed to weathering and probably indicate real difference between the two samples.

3 Diagnostic ratios

One of the improvements of the revised Nordtest methodology is the introduction of diagnostic ratios (i.e. ratios between individual GC-MS peaks or group of peaks) derived from recorded series of alkylated PAHs and generic petroleum biomarkers. To function as a diagnostic ratio it requires that the components from which it is derived describe characteristic and inherent properties of the oil
sample in question. Besides it must be robust, i.e. invariable towards weathering and degradation, and measurable with a low degree of analytical variance [8]. Dalig et al. (these proceedings) have recently described the selection and evaluation of a series of 24 diagnostic ratios (6 PAHs, 18 biomarkers) for the revised Nordtest methodology [2].

For the present oil spill case, we applied a slightly modified series of 30 diagnostic ratios including alkanes (nC17/Pr, nC18/Ph), 12 PAHs and 16 biomarkers; thus, for the PAHs we included additional ratios of some parent PAH against the sum of their alkylated isomers. Relative standard deviations (RSD) of the diagnostic ratios were determined by analysing the spill sample three times. Apart from a couple of ratios derived from small triterpane peaks having rather high RSDs (>15%) all other diagnostic ratios had smaller and rather similar RSDs with an average of 3.1%. This variation is supposed to be related primarily to the analytical instrument and therefore relatively constant within the same series of analysis. The correlation of the diagnostic ratios from the spill and source samples is displayed in Figure 6 (A). The solid line represents 100% correlation and filled circles with error bars the diagnostic ratios.

![Figure 6: Correlation of diagnostic ratios for the recent marine oil spill case (A) and a Nordtest round robin experiment in 2001 (B).](image)

### 4 Statistical evaluation of diagnostic ratios

The revised Nordtest method incorporates a statistical evaluation of the diagnostic ratios for obtaining an objective and technically defensible conclusion of the forensic oil spill identification. A correlation of the diagnostic ratios within 95% confidence limits is the proposed criteria for positive match, and within 98% confidence limits for probable match, between spill and source samples. In the present spill case we observed an average RSD of 3.1%, and for the sake of simplicity we applied this as a general RSD to all diagnostic ratios. Within the 95% confidence limits this corresponds to 8% RSD according to the Student's t-test when analysing triplicate samples [9], and this value was used for error bars in the plot of the diagnostic ratios in Figure 6 (A). For comparison, Figure 6 (B) shows a positive match where all diagnostic ratios of the spill and source samples correlates to 100% within the 95% confidence limits. This example is from a recent Nordtest round robin experiment in which NERI took part in 2001 [10].
For the present oil spill case, the correlation of the diagnostic ratios clearly shows that several ratios (black filled circles) falls outside the 100% line at the 95% confidence limits thereby demonstrating that the two samples do not fulfil the criteria for a positive match. Even within the 98% confidence limits (22% RSD) a couple of ratios fall outside (not shown). As none of the diagnostic ratios seemed to be influenced by weathering, according to the weathering checks, the observed discrepancies in the correlation are probably caused by inherent differences between the samples. According to the revised Nordtest evaluation criteria [2], it is therefore concluded, from a forensic point of view, that the identification of the spill in this case is inconclusive. By the original Nordtest terminology the conclusion would be non-identical.

5 Conclusions

The application of a series of PAH and biomarker diagnostic ratios derived from GC-MS fingerprinting together with statistical evaluation for the forensic identification of a recent marine oil spill displayed minor but inherent differences between the oil spill and the candidate source sample. These differences might easily have passed unnoticed by simple visual inspection and comparison of the chemical fingerprints and become subjective to the analyst’s experience and personal evaluation. The presented case story demonstrates the obvious potential of the revised Nordtest methodology as a forensic tool by rendering the evaluation less subjective and dependent on the analyst’s experience and thereby turning it into a more robust, operational and technically defensible procedure.

References

66  *Oil and Hydrocarbon Spills III*

