Ministerie van Verkeer en Waterstaat

Rijkswaterstaat

Oil Spill identification - Round Robin 2004

The comparison of three gas oil samples

28 February 2005

Ministerie van Verkeer en Waterstaat

Rijkswaterstaat

Oil Spill identification - Round Robin 2004

The comparison of three gas oil samples

28 February 2005

The results of round robin 2004 - Oil Spill Identification between fifteen international laboratories.

Front page:
Photo of the samples in vials, as send to the participants.

Second page:
Photo of the samples as received at the laboratory of RIZA

Colophon

Department	WIL
Information	P. Kienhuis
Telephone:	0031 320 298593
E-mail:	p.kienhuis@riza.rws.minvenw.nl
Authors:	Ing. P. Kienhuis and Dr G. Dahlmann
Client(s)	RIZA dep. WIL and WIK
Dated:	28-February-2005
Status:	Definitive
Document number:	2005.042X

Contents

Abstract

RIZA has organized a round robin concerning oil spill identification between 15 international laboratories
Gas oil samples, originating from a real case, have been send to the participants with the following information:

Enclosed you will find 3 samples numbered I, II and III.
Sample I is from a bunker boat, Sample II is from the surface water and Sample III is from a bunker centre.
Question: Does the oil on the water come from the bunker boat and/ or bunker station?

The spill sample (# II) is compared with both possible source samples (# I and # III).The conclusions are summarized in the table:

	Method		Results	
	FID	MS	Sample I - II	Sample II - III
NFI	-	X	--	--
BSH	X	X	--	--
Sintef-t	X	X	++	--
Sintef-r	X	X	?	--
WGS	X	-	+	+
Cedre	-	X	--	--
WF	X	-	-	+
SKL	X	X	--	--
RIZA	X	X	--	--
WRD	X	-	+	+
WBL	X	-	-	+
LASEM	-	X	--	--
BMM	-	X	++	++
LVA	X	X	--	--
NBI	X	X	++	?
ETC *	X	X	++	--

-- : Significant differences between the samples have been found.

++ : No significant differences between the samples have been found.

? Conclusion is uncertain

- and +: Conclusions are based on GC-FID without GC-MS confirmation.

* The report of ETC was received after the concept report was send to the participants for comments.

Most of the participants that applied GC-MS analyses and compared the ratios of the peaks, especially of m/z 216 (methyl fluoranthenes/pyrenes), found differences and concluded a non-match ("non-identity", etc.) between the spill sample and the two source samples.

Using only the GC-FID method makes it difficult to distinguish between the samples showing quite similar alkane patterns and alkane/isoprenoid ratios.

In chapter 2 of the report the individual participants are mentioned and their original conclusions shown. The complete original reports are published on CD because the information (> 30Mb) is too large.

In chapter 3 the results are shortly discussed. A goal of this round robin is to check the own method and to learn from the methods of other labs. Therefore all original reports are provided and it is strongly advised to the participants to read the reports.
The discussion deals for a large part about the difference in results, when applying two common methods to calculate ratios.

It can be concluded that the round robin has resulted in a very useful set of reports dealing with the comparison of gas oil samples. The individual participants can check and verify their conclusions against a large number of international laboratories.

1. Introduction

1.1 Invitation

The round robin started with this invitation by email with enclosed letter on 13-8-2004:

Dear all,

On behalf of RIZA, I like to invite you to participate in a round robin dealing with oil spill comparison
Information can be found in the invitation.

Letter: Oil spill identification round robin

Dear all,

RIZA (Institute for inland water management and waste water treatment) is a governmental institute responsible for the environmental quality of the surface water in the Netherlands.
One of our tasks is the comparison of oil samples from spills with possible sources.
For many years now we exchange spill samples each year with the Netherlands Forensic Institute (NFI) for quality assurance. The results of this exchange are evaluated in a report with the original reports as annex.
We learned from some other labs, that this exchange of oil samples is quite unique and that they would like to participate in our “round robin”.
Therefore we decided to test this year the willingness of Dutch and international governmental related labs to participate in this round robin.

Samples
Last years samples consist of gas oil, lubrication oil, bilge and bunker oil.
This year our study deals with three gas oil samples, which must be compared. Also a standard oil as reference will be provided.
Every lab has to follow its own standard procedure for oil comparison.
The resulting report should not only consist of a simple yes or no, but also a summary of the procedure should be given, together with detailed results and reasons, why a decision has been taken.
The original reports will be combined in a final report, together with an evaluation of the results.

Time schedule

August:	Request for participation
September:	Delivery of the samples.
October:	Reports have been returned.

November:	Final report will be sent to the participants.

I hope it will be possible for everybody to analyze the samples and make a report within one month, so we are able to finish the whole procedure before the end of this year. If you think you will need more time, please inform me.

Costs
For this year RIZA will pay the costs. Depending on the time needed, it is possible that we need to ask for a contribution next year.

Coded results.
It's our intention to mention participants and their results. It has the big advantage that information can be shared more easily. If this is a reason for you, not to join the round robin please inform me. You may participate anonymously

Participants
This request is sent to Dutch inland water labs, participants of CEN BT 120, participants of Helcom and some known governmental related labs working in this field.
Gerhard Dahlmann (BSH, Germany), who offered to help me with the organization and evaluation, provided the email addresses of the Helcom participants. If you know other governmental related labs that like to join the round robin, please inform me.

If you are willing to participate in this round robin, please send me a confirmation together with a postal address, where to send the samples.

Reactions of four local authorities (Dutch inland water labs), the Netherlands forensic institute (NFI) and nine European laboratories were received.
None of them wanted to participate anonymously. In chapter 2 the participants will be mentioned.

1.2 Samples

At the beginning of September the samples were send to the laboratories with the following letter enclosed:

Subject: Round robin oil comparison: instructions and samples

Dear all,

Enclosed you will find 3 samples numbered I, II and III.
Sample I is from a bunker boat,
Sample II is from the surface water,
Sample III is from a bunker centre.

Samples have been taken in a harbor. Between the bunker boat and the quay next to the bunker centre a thick layer of gas oil was found and it was expected that something has gone wrong during bunkering of the bunker boat. Employees of boat and bunker centre however told that they didn’t cause the spill.

Question: Does the oil on the water come from the bunker boat and/ or bunker station?

For participants who have no experience with gas oil:
About 2 µl sample dissolved in 1 ml solvent and a 1 µl injection gives a good signal on the GC-FID.

The problems associated with the identification of light fuel oil samples are especially well described in http://www.bsh.de/de/Produkte/Buecher/Berichte/Bericht31/Bericht31.pdf. Hints, tips and precautions are given here concerning their analysis and result interpretation.

In my invitation I talked about an oil standard made by Sinteff. In the meanwhile however it became clear that such an action costs much more time (and money) than expected. So we decided not to combine that action with the round robin and therefore you only receive three samples.

Because the method of analysis is free I would like to have a short description of the method, followed by a discussion about the results and where the conclusion is based on. Chromatograms and/ or statistical treatment is very welcome.

In the final report your original contribution will be present as annex. Therefore I have to ask you to send the results in a digital format by email. The main part of the report will be small and give a summary of the results.

I will send you a concept and make it final after receiving your reactions.

1.3 Choices

RIZA is involved in a working group (CEN / BT/Task Force 120 with among others (Sintef, BSH, NERI) trying to produce European guidelines and maybe in future a norm for oil spill identification. The proposed method, based on the revised Nordtest method, has been firstly tested by means of a Round Robin organized by Sintef and with 4 crude oil samples and two HFO samples. A crude oil had to be found out (See Faksness et al).
Because the method is also intended for gas oil, lubrication oil and oil wastes, such as bilge oil and sludge it would be useful to organize Round Robins for these types of oils, too. So we started with gas oil (Light Fuel, Diesel).
In addition, this case was chosen, due to:

- the large amount of spill sample available,
- the very low amount of biomarkers
- Riza has concluded a “non-match” in this case (a severe decision, which was not expected).

1.4 Deliverables

Most results were received by email. The remaining reports were scanned.
A combination of all data would result in a huge report. Therefore the report is published on CD. The individual reports can be found on the CD, whereas in this final report of the Round Robin the information is summarized in chapter 2 and discussed in chapter 3.
The results will not be discussed in great detail. All individual reports are available and all participants can compare their own method with others.

1.5 ETC of Environmental Canada

Zhendi Wang received the samples in December 2004, when Asger Hansen of NERI (Denmark) visited him. The report was received after the concept report was send to the participants for comments.
Because at least one of the participants needs the results of the round robin for certification the contribution of Zhendi Wang is mentioned and discussed separately in Annex A.

2.Individual results

In this chapter a summary of the results of each participant will be given. All original results, separated in directories, can be found on the CD.
Each participant is introduced shortly and the method(s) used to analyze the samples are summarized. The conclusion indicated with the => sign is an exact copy of the conclusion found in the original report.

2.1 Netherlands Forensic Institute (NFI)

Contact: R. de Bruyn

The Netherlands Forensic Institute www.forensischinstituut.nl analyses oil samples in cases of arson (e.g. gasoline and kerosene) and environmental contaminations.
The RR-samples were analyzed with GC-MS. In full scan mode, at a concentration of 1 g/l dissolved in petroleum ether for screening and at 10 g /l for the SIM analysis of PAH's and terpanes. The method uses 14 specifically formulated PAH ratios and 7 terpane ratios, which are separated in two graphs based on stability and weathering behavior according to Venosa et al.
Comparison of the alkane patterns revealed differences.
The applied ratios have been normalized on the sample of the surface water and the result is shown. Several ratios differ too much from the ratios of sample II.

=> Based on the results of the analyses, there is no match between the samples I and II and no match between the samples II and III.

2.2 Bundesamt für Seeschifffahrt und Hydrographie (BSH)

Contact: G. Dahlmann

BSH http://www.bsh.de analyses environmental contamination samples from the North Sea and from the inland of Germany.
The RR-samples have been analyzed with GC-FID and GC-MS in SIM mode. 27 mass-chromatograms were produced for visual comparison. 9 of them, found to be most important for differentiating between oil samples, are read into an Access database called COSI, (developed by Dahlmann). Here 29 compound ratios are produced automatically from every sample. In his report several pictures come from the database. An important advantage of the database is the ability to search for similar samples in all samples analyzed last years, thus evaluating the uniqueness of a sample. A PowerPoint presentation about the database

can be downloaded at BSH.
http://www.bsh.de/de/Meeresdaten/Umweltschutz/Oelidentifizierung/Oeldatenbank.ppt
The color of the samples, alkane patterns and alkane – isoprenoid ratios of the GC-FID analyses and 9 ion ratios of the GC-MS analyses have been use to come to the following conclusion:

=> According to significant differences in the relative concentrations of the compounds, which are detected by mass-fragment 216, a "non-match" between all 3 samples must be concluded. This is supported by the color of the pure samples and smaller differences found in the GC-screening.

2.3 Sintef

Contacts: P. Daling and LG. Faksness

Sintef http://www.sintef.no/ is a large Norwegian engineering institute and is among others active in the petroleum industry and petroleum research. Sintef analyzes all oil spill samples for the Norwegian government.
Sintef has published in 2002 a revised Nordtest method (Nordtest Technical report no. 498) that can be found in the **Daling_Sintef** directory on the CD and on the Internet http://www.nordicinnovation.net/nordtestfiler/tec498.pdf.
Per Daling is convenor of CEN working group 120, which develops two guidelines for respectively sampling (Part 1) and oil spill identification (Part 2). Besides Faksness also Dahlmann and Kienhuis are members of working group Part 2. Points, such as the form of the compound ratios (based on geochemical parameters or not) and the kind of error handling are still controversially discussed in the working group. The latter point is reflected in the report of Sintef.
Sintef analyzed the RR-samples with GC-FID and with GC-MS in SIM mode on PAH's and biomarkers.
The resulting chromatograms are first evaluated visually and when a match is assumed a series of ratios are statistically compared.
In their report a point of discussion in the working group about the statistical evaluation method is mentioned and both methods (RIZA-approach versus Student's t distribution) are applied to compare the samples. In chapter 3 this item will be discussed.

=>No clear conclusions can be drawn based on the RIZA-approach and its recommended criteria, when triplicate analyses are performed. However, based on a total evaluation of this statistical treatment approach of the diagnostic ratios (in tables A.2 – A.4), the following conclusions are suggested:
Sample I: Probable match to the Spill
Sample III: Non-match to the Spill

=> When using the Student's t distribution and criteria as suggested in the revised Nordtest Methodology report no. 498, gives the following conclusions:
Sample I: Positive match to the spill
Sample III: Non-match to the spill

2.4 Waterschap Groot Salland (WGS)

Contact: K. de Haan

Waterschap Groot Salland www.wgs.nl is a local authority responsible for the water quality of an eastern part of the Netherlands.

Local and regional water management in the Netherlands is in the hands of Water Boards. Water Boards are decentralized public authorities with legal tasks and a self-supporting financial system. Water Boards are responsible for flood control, water quantity, water quality and treatment of urban wastewater. Operational task include the management of pumping stations, wastewater treatment plants, maintenance of waterways and flood defense structures. Water Boards are embedded in the general democratic structures. In 1850 there were about 3500 Water Boards. Mergers soon reduced this number. By 1 January 2004 there were 37 Water Boards. Approximately 9,000 people work at the Water Boards.

The analysis of the total amount of mineral oil in water samples (ISO 9377) is a routine method, which is also applied to identify the type of oil in samples and sometimes to compare samples in case of an oil spill. For the analysis GC-FID is used. Resulting chromatograms can be found in the original report.
The results were received by email in Dutch. All information has been translated and combined with chromatograms in the PowerPoint file **Results GrootSalland.ppt**

=> From the analyses it can be concluded, that the oil in all samples consist of diesel (gas oil) and that the distillation pattern and the composition of the hydrocarbons match well. So the 3 samples contain the same oil.
Sample II is coming from the "bunker boat" and/or the "bunker station"

2.5 Centre de Documentation de Recherche et d'Experimentation sur les pollutions accidentelles des eaux (Cedre)

Contact: J. Guyomarch

Cedre (http://www.le-cedre.fr/) was created in 1978 within measures taken after the wreckage of the "Amoco Cadiz", to improve preparedness against accidental water pollution and strengthen the national response organization. Its expertise encompasses both marine and inland waters.

The RR-samples have been analyzed with GC-MS in SIM mode on alkanes, PAH's and biomarkers at two different concentration levels.

=> The comparison of the samples I, II and III led to the following conclusions:
- On the one hand, n-alkanes show differences, especially between samples I+II and sample III in the region n-C11 to n-C24. These variations pointed out also differences between the stored gas oils (samples I and III). On the other hand, the possible evaporation processes affecting these distributions does not allow differentiating sample II from the samples I and III.
Finally, the relative abundances of compounds ranking from n-C25 to n-C32 did not show differences.
- The diagnostic ratios calculated with PAH's commonly analyzed did not present variations that could explain differences of origins. As regards the fragment 216, the analyses carried out in triplicates showed significant differences.
- Finally, the biomarkers allowed differentiating the 3 samples according to the family considered. No replicates were performed for these compounds but the differences are great enough to assume that samples I, II and III have no common origins based on that criterion.

=> The conclusion of this oil spill identification is that the sample collected at sea does not come from the bunker boat or the bunker centre. Moreover, the two bunker gas oils have different origins.

2.6 Wetterskip Fryslân (WF)

Contact: R.B. van der Meer and V. Huremovic.

Wetterskip Fryslân is a local authority (Water Board see 2.4) responsible for the water quality of a northern part of the Netherlands. The analysis of the total amount of mineral oil in water samples (ISO 9377) is a routine method, which is also applied to identify the type of oil in samples and sometimes to compare samples in case of an oil spill.
For the analysis GC-FID is used. Resulting chromatograms can be found in the original report.
The results were received by letter in Dutch. Page 2 contains the most relevant information within the framework of this report and has therefore been translated and can be found on the CD, together with the original report and the chromatograms.

=> All samples consist of diesel oil (C9 to C30)

=>The peak pattern of the diesel oil in the chromatograms of sample 2 and 3 show a very good similarity. Also the ratios of C17/pristane and C18/phytane compare well.

Sample 1 is different however. The C17/prystane and C18/phytane ratios are not similar to the ratios of the other samples.
Based on the analytical results it is concluded, that the oil of sample 2 matches with the oil of sample 3

2.7 Swedish National Laboratory of Forensic Science (SKL)

Contact: M. Källberg

SKL is the only forensic laboratory in Sweden and thus all kinds of investigations are performed, mainly to assist the police. The total number of persons employed is about 240. SKL has four departments:

biology+morphology
documents+IT
drugs
chemistry+technique+weapon.

One group of nine persons is dealing with investigations related to fire, oil and environment. A big part of this is the analyses of oil samples in cases of fire-raising (e.g. gasoline and kerosene) and environmental contaminations.

SKL does not (yet) have a home page on the Internet in English, they only have information in Swedish: http://www.skl.polisen.se/.
The link to the Nordtest method in the report on the CD has been changed since Magnus Källberg wrote the report and the address is now:
http://www.nordicinnovation.net/nordtestfiler/chem001.pdf
This is an older method and not the same as report 498. They do not calculate a lot of ratios and mainly compare the ion chromatograms by eye only. But they have knowledge of the identity of many peaks in the ion chromatograms.

=> Summary of the GC-FID analysis
The findings do not support any oil from the suspected sources, especially not the bunker centre oil (sample III), to be identical with the oil of the oil spill, nevertheless we decided to make a GC-MS analysis.

=> Summary of the GC-MS analysis
There are similarities but also differences between the three oil samples. The most important differences according to our opinion are the following:
The ratio of Ts/Tm is different in the oil spill compared to the suspected sources. The 2- and 3-methylphenantrenes are lower in the water sample than in the suspected sources (washed out?). Dibenzothiophene is lower in sample I than in the water sample (we expected the opposite) and pyrene is lower in the water sample than in the reference oils (pyrene washed out?).

=>As a result of our analysis we conclude that none of the oils from the suspected sources are identical with the spilled oil and express this in

the same way as with other forensic comparisons (please see the last page of the SKL report)
The findings strongly indicate that none of the oil samples I or III were identical with the oil spill sample II at the time of discharge (Level -3).

2.8 Rijksinstituut voor Integraal Zoetwaterbeheer en Afvalwaterbehandeling (RIZA)

Contact: P. Kienhuis

The acronym RIZA stands for Institute for Inland Water Management and Waste Water Treatment. RIZA www.riza.nl is the research and advisory body for the Rijkswaterstaat (the Directorate-General for Public Works and Water Management) for inland water in the Netherlands and a leading international centre of knowledge for integrated water management.
Oil spill identification samples from inland waters and the North Sea are analyzed according the concept guideline Part 2 of CEN BT/TF 120, which is based on the revised Nordtest method.

In general: Samples are analyzed in duplicate with GC-FID. After assessment of the results the (probably) matching samples are also analyzed with GC-MS in SIM mode.
After a visual assessment of the chromatograms (probably) matching samples are integrated for a large range of compounds (depending on the type of oil and compounds available). Ratios between spill and source samples are statistically matched based on the repeatability limit (ISO 5725)

=> On the 5th of June 2004 a contamination of the surface water in Dordrecht has been observed. The analysis of the samples showed that the contamination consists of mineral oil, which can be specified as gas oil. The contamination shows significant differences with the samples of the bunker boat and the bunker centre

2.9 Waterschap Regge en Dinkel (WRD)

Contact: A Kroeskamp

Waterschap Regge en Dinkel www.wrd.nl is also a local authority (Water Board see 2.4) responsible for the water quality of an eastern part of the Netherlands.
The analysis of the total amount of mineral oil in water samples (ISO 9377) is a routine method, which is also applied to identify the type of oil in samples and sometimes to compare samples in case of an oil spill. For the analysis GC-FID is used. Resulting chromatograms can be found in the original report.
The results were received by letter in Dutch. Page 1 and 2 contains the most relevant information within the framework of this report and has therefore been translated. The original document, including the

chromatograms, has been scanned into document **kroeskamp_regge_en_dinkel.pdf** .
The translation of page 1 and 2 can be found in document **Results pag 1 and 2 english.doc** on the CD.

Result:

Sample I and II and III are similar oil samples based on:
- Similar characteristics of the oil chromatograms
- The 2 common fractions (C16-C20 and C20-C24) to compare oil samples show similar percentages.
- Almost the same color (red diesel)
- Similar diesel hump with an alkane series of C10-C30.

Note: The lower C8 peak of vial II is probably caused by evaporation (oil on water results in a large area by spreading of the oil)

2.10 Waterschapsbedrijf Limburg (WBL)

Contact: F. Mertens

Waterschapsbedrijf Limburg www.wbl.nl is a local authority (water board see 2.4) responsible for the water quality of a southern part of the Netherlands. The analysis of the total amount of mineral oil in water samples (ISO 9377) is a routine method, which is also applied to identify the type of oil in samples and sometimes to compare samples in case of an oil spill.
For the analysis GC-FID is used. Resulting chromatograms can be found in the original report.

=> Sample 1 is clearly different from the samples 2 and 3. This chromatogram shows in the area about C 35 a number of components, which are not in samples 2 and 3. These last ones are very much alike.

Conclusion 1: The oil on the water is not coming from the bunker boat.
Conclusion 2: Whether the oil is coming from the bunker centre or elsewhere will be verified with GC-MS analysis, they look very similar.

Unfortunately it was not possible to analyse the samples with GC-MS.

2.11 LASEM

Contact: P. Pierre

LASEM is a laboratory of the ministry of Defense and is located in Brest.

The results were received on paper. The scanned report can be found on the CD.
The samples have been analyzed with GCMS in the full scan mode (m/z 50 – 250).

=> The three gas chromatograms are typical of light fuel oil. In first approximation, they seem to be quite similar.

- Mass 57 - annexes v4 to 9:

The maturation index (CPI) is similar for the 3 gas oils, and all of them are not weathered (WR low);
The normalization on C25 is different for sample III (more alkanes in low boiling range);
The pristane and phytane ratios present small difference;

- Mass 216 - annexes p7-8:

Compounds ratios present significant differences for especially one of them (2Mpyrene/benzo(a)fluorene).

- Masses 220, 234,240, 184 - annexes p10 to 13 :

They do not reveal any significant differences between the 3 samples.

=> The three samples are different fresh light fuel oils.
The oil found on the water does not come from the bunker boat or the bunker center.

2.12 Beheerseenheid Mathematisch Model van de Noordzee (BMM)

Contact: P. Roose

The institute uses three abbreviations depending on the language. See http://www.mumm.ac.be/
In English: The Management Unit of the North Sea Mathematical Models and the Scheldt estuary, abbreviated to MUMM, is a department of the Royal Belgian Institute of Natural Sciences (RBINS), a federal scientific establishment that comes under the Federal Science Policy (previously known as OSTC).

The samples were analyzed with GC-MS in SIM mode according to the standard procedure that is largely based on the revised Nordtest Methodology as described in Faksness et al. (2002) report 498.
The report of MUMM was received by email without the chromatograms in the annexes. The chromatograms were received by post. A relevant selection of 18 of them was scanned and can be found on the CD.

Alkane - isoprenoid ratios and 5 PAH ratios have been compared statistically.

=> The current analysis leads to the conclusion that all samples are from the same source, particularly if the possibility of other sources is absent or highly unlikely. Further more thorough examination of the samples, looking at more DRs and possibly more PAHs could further confirm this.

2.13 LVA

Contact: R. Skolmeistere

LVA (http://www.lva.gov.lv/lea) is a governmental institution subordinated to the Ministry of Environment of the Republic of Latvia. Their aim is to establish a seamless environmental information system to improve the environment and move towards sustainability.

The samples have been analyzed by GC-FID and GC-MS.

1. Gas chromatograms of sample II (from surface water) and sample I (from bunker boat) are different see Fig.1. Peaks in the chromatogram of sample II C10-C14 are higher than the same peaks of sample I.
2. Gas chromatograms of sample II (from surface water) and sample III (from bunker centre) looks similar see Fig.2.
3. Sample II and III weathering check Fig.3 shows that normalized peak areas C11-C12 are higher for sample II, and C16-C17 peak areas are higher for sample III. In normal weathering conditions it is the other way round.
4. Ratios - C17/ pristine, C18/ phytane and pristine/phytane (see table1) don't show convincing differences, but ratios C17/pristine and C18/phytane are higher for sample II.

5. To be sure that sample II is not from bunker boat and bunker center the GC-MS analysis were done. Mass – chromatograms of mass 216 were compared (Fig.4). After normalizing the peak heights on peak 4-M-pyrene (see table2), the sets of five parameters are founded different (for mass 216 of samples II, I and sample III).

=> The sample from the surface water (II) is not identical to the sample from the bunker boat (I) and not identical to the sample from bunker centre (III).

2.14 The National Bureau of Investigation (NBI)

Contact: N. Viitala

The National Bureau of Investigation http://www.poliisi.fi/nbi is one of the national units of the Finnish Police. The NBI is specialized in investigating and preventing serious, organized and international crime. At the national level, the NBI is responsible for criminal intelligence and identification of new types of crime. The NBI also develops techniques for criminal investigation and provides training in the mentioned subjects.

The samples have been analyzed with GC-FID and GC-MS in SIM mode. The chemical analytical methodology is based on the CEN/TC BT WI CSS27002.4 Oil spill identification – Waterborne petroleum and petroleum products- Part 1, Date: 2003-11-31.

=> Based on the visual inspection, the samples 1 and 2 are the same kind of oil. There are some differences between the samples 2 and 3, but they are quite similar oils. The calculated result shows that all these samples are positive matches. We do not have much experience in using statistical methods, nor do we know how much weight we can put on them. After all, the conclusion is that the samples 1 and 2 are positive matches, and the sample 3 is a probable match.

We could not establish from the sample description whether the bunker boat was bumping the oil to or from the bunker centre. We consider that it would have been an important piece of information for the weathering check.

3. Discussion

3.1 Samples

Per Daling wondered, whether the samples are from a real case, because the spill sample was pure and non-weathered. It is correct that most spill samples with gas oil involved consist of a very thin layer on water or are taken with an absorption cloth.
In this specific case a thick layer was present between the bunker ship and the quay at which the ship was moored to. Between the ship and the quay some protection (fenders or bumpers) was available, which also hindered the oil to spread. Unfortunately no photos were taken, mainly because the samples were taken at night. During sampling the opening of the bottle was kept just below the water level in order to sample as much oil as possible.
RIZA concluded that the samples are different, what sounds very strange when looking at the situation. Discussing this with the sample taker revealed that more tanks were present at the bunker station and that it could be possible, that the wrong tank(s) have been sampled.

3.2 Methods

Dahlmann(BSH) and Kroeskamp(WRD) mentioned, that the color of the samples are not equal. From the photo in the report of Dahlmann and the photo of the front page it can be seen that sample III is somewhat different from the rest. The bottles, as received, can be seen on page 2 (sample III is left). Here the difference is less obvious, probably because of the thicker layer.

An injection concentration advice for the GC-FID analysis of 2 ul gas oil/ml solvent was given in the letter belonging to the samples. Most of the participants used about this concentration for GC-FID and GC-MS but to be able to see more of the biomarkers a 10 times or higher concentration was used by several participants with Guyomarch (Cedre) on top with 0.8 ml gas oil in 0.8 ml DCM. Some of the results show good terpane and sterane chromatograms, while most of the participants didn't use the biomarkers for comparison at all.
RIZA has never tested the use of biomarkers for gas oil. A large amount of sample has to be injected on the column, which probably influences the properties of the column in separation and resolution. In addition, when ratios are produced, small peaks on a hump have to be measured.

Note:
If some of the participants have statistical information about the repeatability of the analysis of biomarkers in gas oil, it would be useful for all of us to share this information

The Dutch Water Board laboratories have all used the ISO 9377 method to compare the samples. This is a standard method to analyze quantitatively the total amount of oil between C10 and C40 in the chromatogram. The method, often used in combination with large

volume injection is not intended to get a good separation, because only the total amount is important. The method can more or less be used to identify the type of oil, but is obviously not suitable to differentiate between almost identical samples.
Most of the participants, that used GC-FID for screening, were able to conclude that the samples were different based on the alkane pattern. The alkane/isoprenoid ratios have been used by almost all participants but couldn't really help to differentiate the samples because the ratios are almost the same.
Because it was expected that several participant had no experience with gas oil spill samples we advised to read the features of Gerhard Dahlmann. Based on the hundreds of samples in his database he found that most of the PAH concentrations are quite similar. For the three samples of RR2004 Pierre (Lasem) also concluded this for m/z 220, m/z 234, m/z 240 and m/z 184. M/z 216, showing the methylfluorantenes and methyl pyrenes, however is an exception. The retention times of the m/z 216 peaks are higher than C20. After a weathering check it can often be used for comparison.
Eight of the fourteen participants used some or all of the peaks of this ion chromatogram for ratio calculation and comparison.

RIZA has found m/z 230 (i.e. m/z 216 + 14) also to be suitable for gas oil identification.
Recently I combined the m/z 230 (dimethyl fluoranthenes and dimethyl pyrenes) chromatograms of all cases of 2004 dealing with gas oil and gas oil / lubricating oil and found also differences in the patterns of this mass between the samples, while samples which were concluded to be the same also showed an identical m/z 230 pattern.

3.3 Parameter calculation

A point of attention is the use of the two types of ratios (A/B and 100 * A/(A+B)) to compare the samples. Most of the participants applying GC-MS analyses also use ratios to compare the samples. In older methods e.g. chapter 30 of the Bonn agreement of 1993 based on the previous Nordtest method, only a visual comparison is used.
In the concept guideline of CEN /BT/TF 120 it is strongly advised to use ratios to make the comparison less dependent of individual experience. Based on the relative standard deviation (RSD) of the ratios caused by the analytical method, ratios can be compared statistically, i.e. the significance of a given difference is tested. Both ratio formulas are mentioned in the guideline and the working group agreed that an RSD of 5% is reasonable.
The results of RR2004 however revealed that these methods can result in different conclusions. To evaluate this, RIZA has used the validation measurements performed recently to test the suitability of a newer GC-MS for oil spill identification. The spreadsheet with the results of eight standard analyses (Standard alkanes for GC-FID and Standard Brent for GC-MS) can be found in the RIZA directory and is called **Validation MS8v2.xls**.

The results based on the same data, but calculated with the two ratio formulas are summarized in Table 3.1

Compounds	mean	stdev	RSD in %
DMN1 + DMN2	62882847	5410272	8.6
benzo(b)-fluorene	439836	33714	7.7
4-M-pyrene	698076	56734	8.1
Ts	562480	47256	8.4
Tm	446822	36681	8.2
PENT6 (30ab)	3287978	263827	8.0
Ratios A/B			
DMNs/30ab	19.12	0.209	1.1
BbF/4Mp	0.63	0.023	3.7
Ts/Tm	1.26	0.011	0.9
ratios A/A+B			
DMNs/30ab	0.95	0.001	0.1
BbF/4Mp	0.39	0.009	2.3
Ts/Tm	0.56	0.002	0.4

Table 3.1
The st dev and RSD of some ratios calculated with two method and based on eight analyses.

The areas of the individual compounds have a relative high RSD of 8% due to a small decrease of the sensitivity during the series. Using ratios however compensates for sensitivity differences as can been seen from the results. It is also visible that the RSD is different depending on the type of ratio used. The RSD% is much smaller, when ratios are used of the form A/(A+B).

To study the effect further also an additional sheet with eight theoretical values has been added to **Validation MS8v2.xls**.

Compounds	mean	stdev	RSD in %
C1	1.000	0.000	0.000
C2	0.096	0.014	14.627
C3	0.963	0.141	14.627
C4	9.625	1.408	14.627
Ratios A/B			
C2/C1	0.096	0.014	14.627
C3/C1	0.963	0.141	14.627
C4/C1	9.625	1.408	14.627
ratios A/A+B			
C2/(C2+C1)	0.088	0.012	13.306
C3/(C3+C1)	0.488	0.036	7.375
C4/(C4+C1)	0.904	0.012	1.363

Table 3.2
The st dev and RSD of some ratios calculated with two method and based on eight theoretical values.

C1 is the reference peak and has a value of 1. This situation can be reached in real samples by normalizing on the reference peak.

The eight C2 concentrations are randomly chosen and vary around 0.1. The individual C3 concentrations are the same as C2, but multiplied by 10 and the C4 by multiplying with 100. It results in the same RSD for the individual values of C2, C3 and C4. The ratio A/B results in the same RSD%, but the ratio A/(A+B) is dependent on the value of the ratio, which makes it more difficult to use for comparison. The ratio

A/(A+B) tends to A/B when A becomes small and A/A when B becomes small.

It is possible to convert values given in the form 100%A/(A+B) into the form A/B (and vice versa) by simply transforming the formula. For producing the following diagram, 9 pairs of 100%A/(A+B) values were produced (yellow and pink points), each having a difference of 10% around their individual means for mean values of 10, 20, 30 etc. (blue points).
Then the values were converted into A/B-values and the %differences of these pairs around their means were calculated (red crosses, red line).

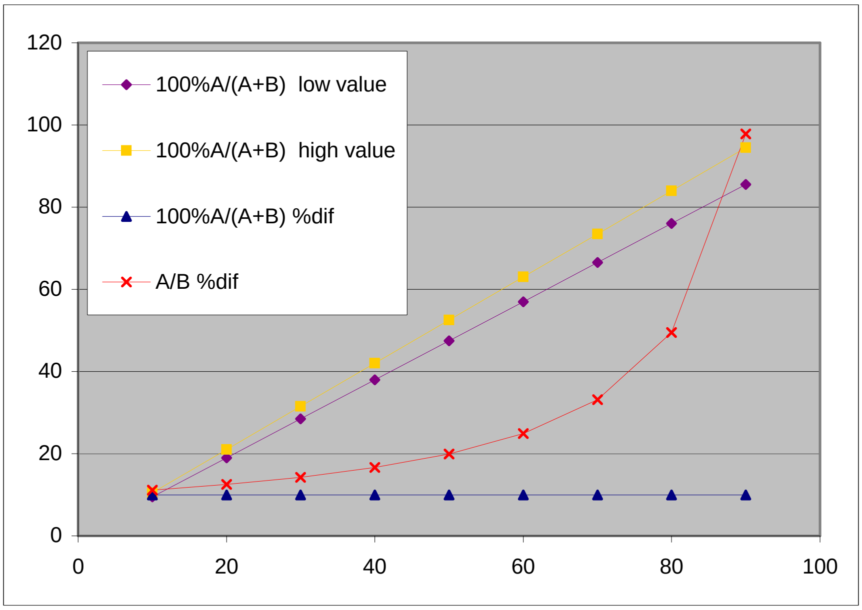


Fig 3.1
10% repeatability (blue points) of two A/(A+B) values, around 10, 20 etc. (yellow and pink points), compared to % difference, when these numbers are converted to the % difference of A/B (red points)

The formula 100*A/(A+B) obviously always produces (in the region of higher values much) smaller %differences.
Differences of A/B-values of 20% (around the value of 50, an exact value of 50 means that A = B), detectable in the mass chromatograms by the human eye, are calculated to be only 10%, when the formula 100*A/(A+B) is used. This effect increases, when B becomes smaller than A, and merely 10% are even calculated for nearly 100% peak differences in the mass chromatograms (when B is very small compared to A).

=>A discussion in the CEN-working group how to deal with this behavior is strongly needed.

3.4 Results

Note:
E.g. Kallberg (SKL) and Dahlmann appointed that the question: Does the oil on the water come from the bunk boat and/ or bunker station?
is wrong. Only a match or non-match can be concluded.
It is indeed true that my question was not well formulated. In all cases only a judge can make this decision based on all evidence.

The results described in chapter 2 and annex 6.1 are summarized in Table 3.3.

	Method		ISO	m/z	PAH &		Results	
	FID	MS	9377	216	bio-markers		I-II	II-III
NFI	-	1		-	A/B		NM	NM
BSH	2	2		A/B	A/B		NM	NM
Sintef-t	1	3		A/(A+B)	A/(A+B)		M	NM
Sintef-r	1	3		A/(A+B)	A/(A+B)		PM	NM
WGS	1	-	+	-	-		M	M
Cedre	-	3		A/B	A/(A+B)		NM	NM
WF	1	-	+	-	-		NM	M
SKL	2	1		-	A/B		NM	NM
RIZA	2	2		A/B	-		NM	NM
WRD	1	-	+	-	-		M	M
WBL	1	-	+				NM	M
LASEM	-	3		A/B	-		NM	NM
BMM	-	1/3		-	A/(A+B)		M	M
LVA	1	1		A/B	-		NM	NM
NBI	3	1/3		A/(A+B)	A/(A+B)		M	PM
ETC*	1/3	1/3		-	A/B		M	NM
Method: number of analyses /sample/method								
m/z 216: Is the ion m/z 216 used statistically and which ratio type is used								
Biomarkers: Are biomarkers used statistically and which ratio type is used								
Results: NM : non-match PB:probably match M:match								
* The report of ETC was received after the concept report was published.								

Table 3.3 Results
A summary of the analysis methods and the individual results.

Sintef has used two methods to compare the MS ratios. The results of both the Student's t-test (Sintef-t) and the repeatability test (Sintef-r) are taken into account.
Table 3.4. gives a summary after removing the results based on ISO 9377.

	All data				Without ISO 9377		
	M	PM	NM		M	PM	NM
I-II	6	1	9		3	1	8
II-III	5	1	10		1	1	10

Table 3.4 Conclusions
A summary of the conclusions

It can be concluded that most of the participants found differences between the samples and reported a non-match for both source samples. The remaining results will be discussed (See also annex 6.1).

3.4.1. Sintef

Sintef used the A/(A+B) ratio method and compared two statistical evaluation methods.
The effect of the applied ratio on the RIZA data is shown in Table 3.5 and Table 3.6

	Sample comparison					
	mean sample 1 bunker boat	mean sample 2 water	mean	absolute difference	repeatability limit based on RSD 3.536% **10.00**	flag
BaF/4Mp	1.658	2.136	1.897	0.478	0.190	1
BbF/4Mp	0.391	0.565	0.478	0.174	0.048	1
2Mp/4Mp	0.632	0.622	0.627	0.010	0.063	
1Mp/4Mp	0.743	0.726	0.735	0.017	0.073	

Table 3.5
RIZA data comparing sample I and II based on the A/B ratio

For BaF/4Mp the absolute difference is 0.478 and relative to the mean 25,2%.

	Sample comparison					
	mean sample 1 bunker boat	mean sample 2 water	mean	absolute difference	repeatability limit based on RSD 3.536% **10.00**	flag
BaF/(BaF+4Mp)	0.624	0.681	0.652	0.057	0.065	
BbF/(BbF+4Mp)	0.2809	0.3609	0.321	0.080	0.032	1
2Mp/(2Mp+4Mp)	0.387	0.383	0.385	0.004	0.039	
1Mp/(1Mp+4Mp)	0.426	0.420	0.423	0.006	0.042	

Table 3.6
RIZA data based on the A/(A+B) ratio

For BaF/(BaF+4Mp) the mean is 0.652 (or 65%; see note) and the absolute difference is 0.057 (or 6.6%) and relative to the mean 8.7%. Sintef reports in Table A.2 of their report a mean of 60% and a difference of 6.6% resulting in 11%relative to the mean. So the results are almost equal and if RIZA had used the A/(A+B) a match would have been found for that specific ratio.
In Table 3.6 only the ratio BbF/(BbF+4Mp) shows a clear difference, but Sintef didn't use this ratio.

Note
The actual ratio Sintef used is A/(A+B)*100% resulting in a % between 0 and 100

So, if Sintef had used the A/B ratio, this would have resulted in a difference of 27% around the mean of the BaF value, a difference that highly exceeds laboratory precision limits.
In the table of results given by Sintef, there are at least two other pairs of values for sample 1 and sample 2, which would have revealed bigger differences, if parameters of the form A/B were used:
the %difference of Ts would have been calculated to be 22% and the TA1 difference even to be 31% (For TA1 the 100*A/(A+B) values given were 94,2 and 92,2 for samples 1 and 2 respectively, which means a %difference of the mean of these values to be as small as 2.1%!).
Although the biomarker calculations must be taken with caution in this gas oil case (see remarks above), Sintef must have come from positive match to non-match, if values were used in the form A/B.

Even the 98% limit, calculated by using Student´s t to be 20.1% (see below), would have been exceeded for at least 3 parameters.

Sintef compared two statistical methods. The Student's t-test original described in the revised Nordtest method starts with a triplicate analysis of one of the samples, preferable the spill sample. Based on the standard deviation of the ratios, the t-test is used to calculate the acceptable difference between the ratios for each ratio. A weak point of the method is that the standard deviation calculated from three measurements is not very robust. Sometimes very low acceptable differences were found resulting in false negative results. Based on the experience of several labs the standard deviation was therefore fixed on 5%. This value is by Sintef used for both evaluation methods.
The use of the repeatability limit (r95%) in the RIZA proposal is based on ISO 5725. Using the standard deviation of 5% a difference ($r_{95\%}$) of 5*2,83 is ~14% is acceptable. This is somewhat higher than the value of 12.4% from the t-test mentioned in the report of Sintef. The value of 14% has been used in the RIZA report to test the difference of two analyses of the same sample. The combination of two or more analyses into a mean value reduces the variance of the mean with $\sqrt{n}$.
Therefore an acceptable difference of 10% is used to compare the mean of the duplicate analyses of each sample. Sintef has analyzed all samples in triplicate so a difference of 8.1 % for the repeatability limit should be used for comparison. This is done for the repeatability limit method, but not for the t-test method. Here it stayed on 12,4%.

Generally, the Students t-test is used to estimate the variance of the mean of a limited number of measurements. Murphy and Morrison give a good example of the use of the t-test.

3.4.2. WBL

Mertens (WBL) reports that the chromatogram of sample I shows extra peaks at the end of the chromatogram.

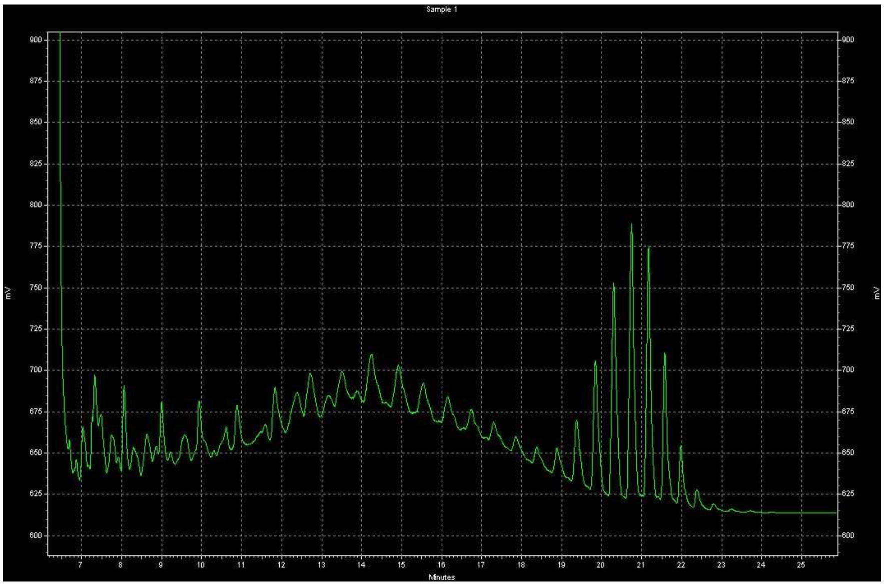


Fig 3.1
Chromatogram of sample I of WBL

The rest of the participants didn't mention this. It looks like a paraffin sample mixed up with sample I. It could be a contamination of the sample or ghost peaks from the previous injection.

3.4.3. MUMM

Roose analyzed the samples with GC-MS on alkanes, PAH's and biomarkers. For comparison the alkane/isoprenoid ratios have been used. These however are quite comparable. The alkane pattern is not compared.
Ratios are used to compare several PAH ratios. As already remarked by other participants only small differences were found for those compounds. Peaks of m/z 216 are however not used.
Using the alkane pattern and the m/z 216 peaks would probably have lead to a non-match for both samples.

3.4.4. NBI

Viitala used GC-FID and GC-MS to compare the samples. Different from the others split injections have been used resulting in good chromatograms and good biomarker patterns.
Evaluation of the alkane pattern revealed differences but it was not clear, whether sample I or III could have been weathered. On this point the information in the second letter, accompanying the samples, was probably not clear enough.

It is obvious that the values of the hopane series cannot be correct, except probably the Ts value. Obviously something has happened with the data.
May be, this would have easier be noticed, if also here the formula A/B would have been used. Values of around 80 up to 97 mean that the relation of the peaks of the hopane series in the mass chromatograms compared to hopane itself is 4:1 up to 32:1, i.e. values, which are not reasonable.
Generally, it should be possible to check the produced parameters by looking at the mass chromatograms. In this sense, the 100%*A/(A+B) values are much more “nebulous”.

The evaluation of the m/z 216 peaks shows clear differences in the report of NBI for at least BbF4MP, but inexperience with the statistical comparison method (as mentioned in the text) has resulted in a match. It can be concluded that the analyses have been performed well, but that emphasis is needed for the statistical evaluation method.

3.4.5. ETC

See Annex 6.1

4. Conclusion

15 laboratories have compared three gas oil samples.

Using only the GC-FID method (4 participants) makes it difficult to distinguish between the samples showing quite similar alkane patterns and alkane/isoprenoid ratios.
May be, this Round Robin can be used a little bit for argumentation that a GC/MS-instrument is definitely needed in oil spill identification.

Compound relations were produced by most participants that applied GC-MS analyses, and the analytical error was used in order to find out whether the observed differences are significant or not.
Most of them found significant differences and concluded a non-match between the spill sample and the two source samples (7 out of 11).

It has been shown that the conclusions drawn by the residual 4 participants, which used GC/MS and which did not find significant differences (and concluded "positive match") were "explainable" (see chapter 3.4 and 6.1).

So, the conclusion of RIZA in this special case of gas oil samples, i.e. "non-match" of all three samples, is highly confirmed (I have to apologize that I have used this case for a Round Robin because in this case the actual "truth" is not known. To be honest, I wanted to hear your opinion).

This Round Robin was especially important because:
Evaluating the use of two formulas to calculate the ratios (A/(A+B) and A/B) revealed that the ratio A/(A+B) has a ratio dependent standard deviation. When comparing two samples, the %difference of the mean of these values is always (partly much) smaller than in case A/B is used. Since oil identification means oil discrimination, i.e. differences between oil samples have to be found out and these differences have to be tested for their significance, using 100*A/(A+B) does not seem to be an appropriate way for calculating parameters.

5. References

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

"CEN/BT/Task Force 120 Oil Spill Identification": Part 2 Analytical methodology and interpretation of results" (TC BT WI 27002, 2004-10, called CEN-part 2 guideline).

Introduction to Environmental Forensics, B.L. Murphy and R.D. Morrison. pag 391-428, 2002 Academic Press ISBN 0-12-511355-2

Harris, D.C. (1995) Quantitative chemical analysis, 4th edition. W. H. Freemantle and Company, New York, USA

Faksness, L.G., H. Weiss, P. S. Daling (2002). Revision of the Nordtest Methodology for Oil Spill Identification – Technical report. Nordtest Technical Report no. 498. (SINTEF report STF66 A01028, 2002).

Faksness, L.G., P.S. Daling, A.G. Hansen (2002). Round Robin Study - Oil Spill Identification. Environmental Forensics 3 (2002)p279-291.

Wang, Z, Fingas M.F. Development of oil hydrocarbon fingerprinting and identification techniques. Marine Pollution Bulletin 47 (2003) p 423-452

6. Annexes

6.1 Emergencies Science and Technology Division (ESTD), Environmental Technology Center (ETC) of Environment Canada.

Contact: Zhendi Wang

The EMERGENCIES SCIENCE & TECHNOLOGY DIVISION of Environment Canada carries out research and development on the properties, behaviour, effects, and in-situ treatment of spilled hazardous materials. Techniques are also developed, and applied, for measuring contamination in air, water, and soil at spill sites and for airborne remote sensing of spills.
http://www.etc-cte.ec.gc.ca/organization/estd_e.html

Zhendi Wang did not receive the samples from RIZA. Asger Hansen and Jan Christensen of NERI (Denmark) have given him a part of their samples when they visited ETC in December 2004.
The report of Zhendi was received after the concept report has been send to participants for comment. Therefore the report of ETC is discussed separately.

Note
NERI has received samples, but was not able to send in a report on time and is therefore not further mentioned in the report.

ETC applies a quantitative method to analyze and compare the samples. The description of the sample preparation method is rather complex and therefore summarized in Fig 6.1

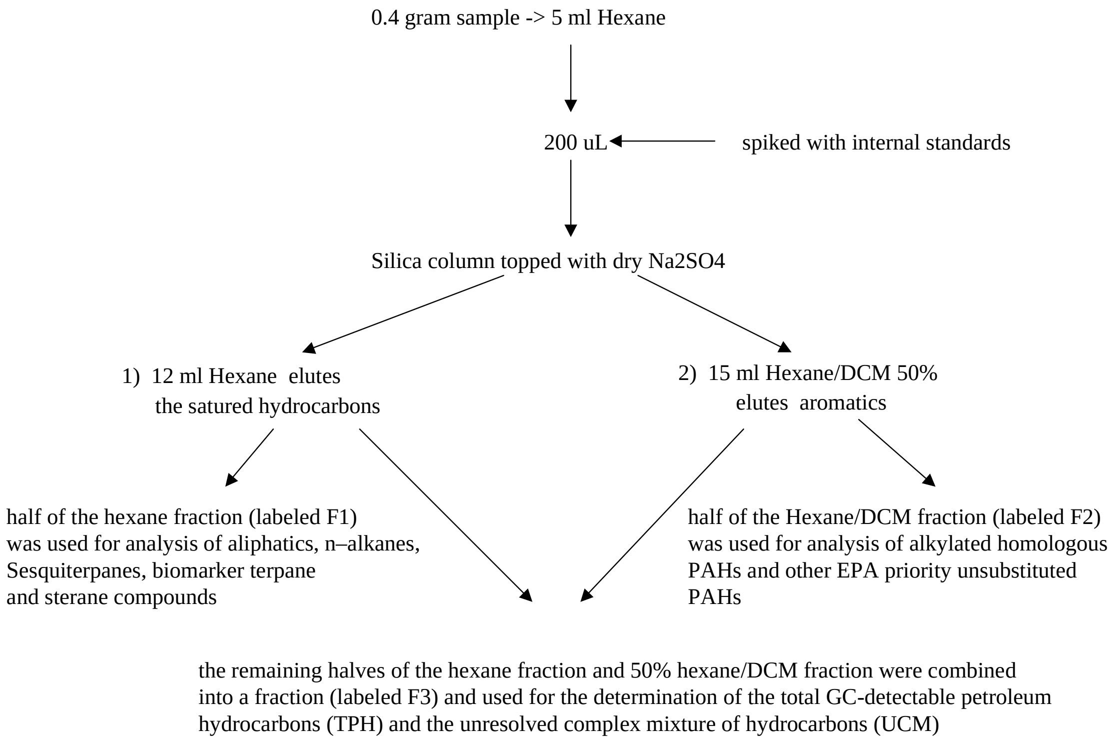


Fig 6.1
Scheme of the preparation of the samples

Samples are analyzed both with GC-FID and GC-MS. Additional to the quite common alkanes, PAH's and biomarkers also the sesquiterpanes (see Wang et al) have been analyzed.

=> (1) Three oil samples are diesel type fuel. They show very similar GC-FID chromatographic features.
=> (2) Chemical fingerprinting results, in particular the characterization results of sesquiterpanes and PAH compounds, indicate that the spilled sample 2 come from the sample 1. That is, the spilled oil found on the water surface between the bunker boat and the quay next to the bunker center was from the Bunker Boat.
=> (3) Sample 3 demonstrates different fingerprints of sesquiterpanes and target PAH compounds. The characterization results of high molecular weight terpane and sterane biomarkers further confirm this conclusion.
=> (4) this work demonstrates that, fingerprinting highly abundant sesquiterpanes in lighter refined petroleum products can provide another very useful means for spill source identification, in particular for those lighter refined product samples which have very similar bulk chemical compositions and therefore difficult to be identified by most GC fingerprinting techniques.

Discussion

The report describes an impressive method to analyze oil samples and to calculate the amount of specific fractions of the samples.

A lot of ratios are calculated by using the ratio formula A/B. To compare samples in cases of oil spill identification the chromatograms and ratios are visually compared.
The use of statistics to estimate the variance of the analysis method and to calculate limits is not mentioned. Several internal standards are added during cleanup of the samples and used to calculate absolute concentrations of compounds and compound groups. It would be interesting how the variance of the absolute values between samples relate to the variances of ratios in single samples.

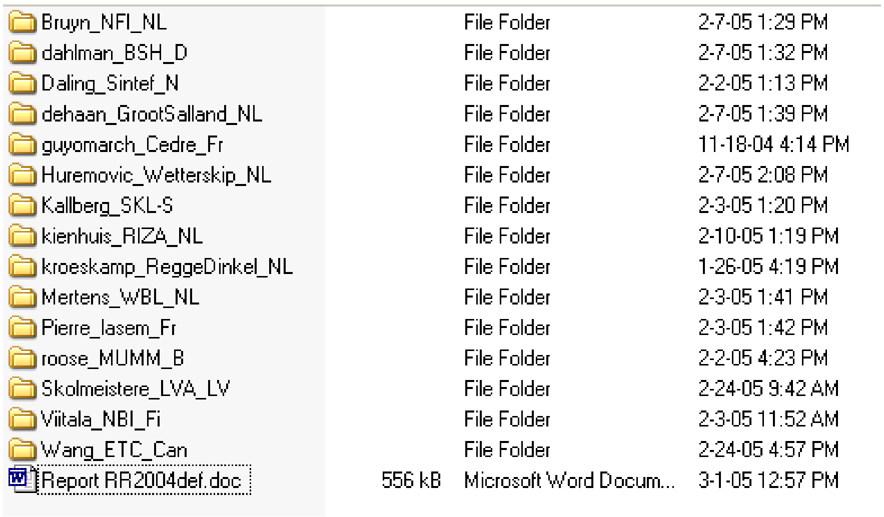
ETC concluded a non-match between the spill sample (II) and the source sample from the bunker centre (III). This is based on differences found in the total PAH concentration, sesquiterpanes and biomarkers.

ETC concluded a match between the spill sample (II) and the source sample from the bunker boat (I). Most of the participants found differences in the pattern of the lighter alkanes and the compounds visible in the ion chromatogram of m/z 216.
Although analyzed, integrated (Table 2) and shown in a graph (Fig 3) the alkanes are not further compared in the ETC report. The results of the GC-FID analyses shown in Fig 1 are difficult to compare, but the MS results shown in Fig 2 show a clear difference in the peak height of C15 relative to C16. This is however not reflected in the concentrations mentioned in Table 2.

The sesquiterpanes are mentioned in the review of Wang et al (2003) as an interesting group to differentiate between light oil samples. The compounds are much more resistant against biodegradation than alkanes and differ in pattern between samples.

6.2 Directory information of the CD

This report is published on CD. The annexes with the individual reports can be found on the CD in directories.



Bruyn_NFI_NL		File Folder	2-7-05 1:29 PM
dahlman_BSH_D		File Folder	2-7-05 1:32 PM
Daling_Sintef_N		File Folder	2-2-05 1:13 PM
dehaan_GrootSalland_NL		File Folder	2-7-05 1:39 PM
guyomarch_Cedre_Fr		File Folder	11-18-04 4:14 PM
Huremovic_Wetterskip_NL		File Folder	2-7-05 2:08 PM
Kallberg_SKL-S		File Folder	2-3-05 1:20 PM
kienhuis_RIZA_NL		File Folder	2-10-05 1:19 PM
kroeskamp_ReggeDinkel_NL		File Folder	1-26-05 4:19 PM
Mertens_WBL_NL		File Folder	2-3-05 1:41 PM
Pierre_lasem_Fr		File Folder	2-3-05 1:42 PM
roose_MUMM_B		File Folder	2-2-05 4:23 PM
Skolmeistere_LVA_LV		File Folder	2-24-05 9:42 AM
Viitala_NBI_Fi		File Folder	2-3-05 11:52 AM
Wang_ETC_Can		File Folder	2-24-05 4:57 PM
Report RR2004def.doc	556 kB	Microsoft Word Docum...	3-1-05 12:57 PM

Fig A1
Directory information of the CD