Technical summary

www.environment-agency.gov.uk

Use of Pattern Recognition to Identify the Source of an Oil Spill on an Inland Water

R&D Technical Summary E1-109/TS

Gas Chromatography (GC) with Flame Ionisation detection (FID) is internationally recognised as the standard approach in the analysis of oil spill pollution incidents. In this technique oil is characterised by temperature programmed capillary chromatography. The components are separated approximately in order of boiling points and displayed as a pattern of peaks (profile) or signature of the hydrocarbon composition of a sample. The distinctive features of these patterns are used to characterise the oil.

Oil products released into the environment undergo weathering processes with time. These processes include evaporation, leaching, chemical oxidation and microbial degradation. Because these processes affect the pattern of peaks obtained, an understanding of weathering effects is valuable to all environmental laboratories.

The protocol used by the Environment Agency's National Laboratory Service for the determination and comparison of oils is predominately done by eye. Although data generated by this method has been successfully used as evidence to support legal cases against suspected polluters, it does have some shortcomings and is seen as subjective.

A feasibility study was initiated to determine whether the use of techniques of pattern recognition could be used to remove the subjectivity of visual methods and support the analyst's findings, increasing courtroom credibility. The R&D Technical Report E72 concluded that the use of sophisticated pattern recognition techniques could be used for the determination of oil-type and oil-spill source in pollution incidents. A phased programme of research leading to the development of an automated software package that would improve the defensibility of data used as evidence in court was recommended:

Stage 1: carry out a detailed feasibility study for a proposed oil-spill source identification system.

Stage 2: develop and test an automated system for the determination of the type of refined oil present in a sample.

Collaborative R&D project E1-109 had three stages, the overall objectives of these stages were:

1. To produce an automated classification system.

2. To produce an automated oil-spillage source identification system based upon the determination of the oil-spillage type.

3. To produce a system which can be used in the Agency's laboratories and to provide evidence for use in litigation.

The initial objective, automated oil-type identification, was developed using Microsoft Visual Basic 6 and is operated through windows, making it user friendly. The software is interfaced to a database of standard oil-type templates, which can be independently updated to include other oil type information. The oil type is characterised by a set of *n* principle peaks (paired heights and retention times), where n may vary for different oil types, together with a number of weathering functions for each peak corresponding to different weathering conditions. The sets of oil type principal peaks stored in the database are referred to as templates.

A sample GC is selected and automatically compared to each of the templates from the database of standard oil types. The template peaks are linked to suggested matching peaks from the sample and an indication of the match quality is given (based on the sum of Euclidian distances between template peaks and corresponding matched peaks from the sample). A manual option to *stretch* and *translate* the sample GC allows for adjustments so that a better match can be attempted, such in the case of retention time shift.

ENVIRONMENT AGENCY

Extending the methodology used for the oil classification has facilitated development of the oil-spill identification software. The sample GC is treated in much the same way as a template, except that a full range of peaks is used rather than selected principal peaks. Each of the potential sources is compared with the sample using the same procedure as that developed for type classification. Because the sample has already been classified, and each of the potential sources is assumed to be of the same type, each GC trace can be easily calibrated so that it is directly comparable with the original type template from the classification database.

In tests the type-matching component successfully matched 99 out of 100 test sample files, and the source-matching component matched 47 out of 52 field sample files.

The project team at Staffordshire University have provided the Agency with a high quality AI "state of the art" software program that has fulfilled the specifications outlined in collaborative R&D project E1-109.

The National Laboratory Service (NLS) is committed to the development of alternative analytical methods, in particular those with greater efficiencies in time, cost, safety and environmental impact.

This R&D Technical Summary relates to information from R&D Project E1-109 reported in detail in the following output:

R&D Technical Report E1-109/TR
Use of Pattern Recognition to Identify the Source of an Oil Spill on an Inland Water
ISBN Number: 1844322653 **November 2003**

Internal Status: Released to Regions
External Status: Publicly Available

Project Manager
Wayne Civil National Laboratory Service, Starcross

Research Contractor
Centre for Intelligent Environmental Systems
Staffordshire University

Copies of these documents are available internally from Regional Agency Libraries or the National Information Centre in Bristol, and externally from the Environment Agency's R&D Dissemination Centre, c/o WRc Information Resources, Frankland Road, Blagrove, Swindon, Wiltshire SN5 8YF, Tel: 01793 865138, Fax: 01793 514562. Website URL: www.eareports.com

Rio House
Waterside Drive
Aztec West
Almondsbury
Bristol
BS32 4UD

Tel: 01454 624400
Fax: 01454 624409