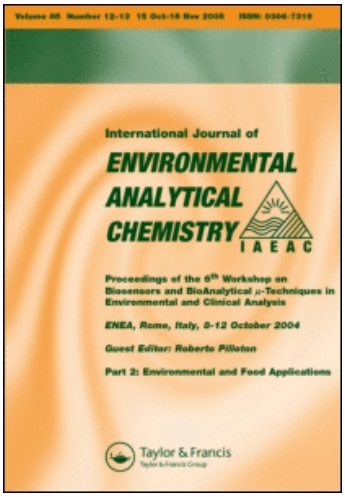


International Journal of Environmental Analytical Chemistry
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713640455

A comparison of direct thermal desorption with solvent extraction for gas chromatography-mass spectrometry analysis of semivolatile organic compounds in diesel particulate matter

Lisa A. Graham[a]; Anthony Tong; Gary Poole[a]; Luyi Ding[a]; Fu Ke[a]; Daniel Wang[a]; Gianni Caravaggio[b]; Jean-Pierre Charland[b]; Pamela MacDonald[b]; Ajae Hall[b]; Yu Cheng[c]; Jeffrey R. Brook[c]
[a] Environment Canada, Environmental Science and Technology Centre, Ottawa, ON, Canada K1A 0H3 [b] Natural Resources Canada, CANMET Energy Technology Centre Ottawa, Ottawa, ON, Canada K1A 1M1 [c] Environment Canada, Air Quality Research Division, Downsview, ON, Canada M3H 5T4

Online publication date: 24 March 2010

To cite this Article Graham, Lisa A. , Tong, Anthony , Poole, Gary , Ding, Luyi , Ke, Fu , Wang, Daniel , Caravaggio, Gianni , Charland, Jean-Pierre , MacDonald, Pamela , Hall, Ajae , Cheng, Yu and Brook, Jeffrey R.(2010) 'A comparison of direct thermal desorption with solvent extraction for gas chromatography-mass spectrometry analysis of semivolatile organic compounds in diesel particulate matter', International Journal of Environmental Analytical Chemistry, 90: 7, 511 — 534

To link to this Article: DOI: 10.1080/03067310903108352

URL: http://dx.doi.org/10.1080/03067310903108352

Intern. J. Environ. Anal. Chem.
Vol. 90, No. 7, 15 June 2010, 511–534

Taylor & Francis
Taylor & Francis Group

A comparison of direct thermal desorption with solvent extraction for gas chromatography-mass spectrometry analysis of semivolatile organic compounds in diesel particulate matter

Lisa A. Graham[a]*†, Anthony Tong[a]‡, Gary Poole[a], Luyi Ding[a], Fu Ke[a], Daniel Wang[a], Gianni Caravaggio[b], Jean-Pierre Charland[b], Pamela MacDonald[b], Ajae Hall[b], Yu Cheng[c] and Jeffrey R. Brook[c]

[a]*Environment Canada, Environmental Science and Technology Centre, 335 River Road, Ottawa, ON, Canada K1A 0H3;* [b]*Natural Resources Canada, CANMET Energy Technology Centre Ottawa, 1 Haanel Drive, Ottawa, ON, Canada K1A 1M1;* [c]*Environment Canada, Air Quality Research Division, 4905 Dufferin Street, Downsview, ON, Canada M3H 5T4*

(*Received 5 November 2008; final version received 9 June 2009*)

Direct thermal desorption-gas chromatography-mass spectrometry (DTD-GC-MS) is a technique that is finding application in the characterisation of the semivolatile organic carbon fraction of ambient and combustion source particulate matter (PM) collected on filters. In this study, three DTD-GC-MS methods were assessed and compared to a conventional solvent extraction method for analysis of a mixture of target analytes in solution and of diesel PM collected on quartz filters. The target analytes included *n*-alkanes, hopanes, steranes and polycyclic aromatic hydrocarbons. This study showed that while the three DTD-GC-MS methods were generally comparable to the solvent extraction method, (1) the choice of calibration strategy and calibration materials has a significant impact on the measured accuracy of a method; (2) very large variations were seen in all methods for the more volatile compounds such as C_{10} to C_{13} *n*-alkanes and naphthalene; (3) accuracy, defined as difference from the known concentration of a liquid sample, ranged from 5% to 32%; (4) precision, defined as the relative standard deviation, ranged from 4% to 16%. The average difference of DTD-GC-MS results from the solvent extraction results for the diesel PM filters ranged from 20% to 40%. This difference was driven by the large number of target analytes present at relatively low concentrations ($<25\,pg/mm^2$) and their corresponding higher variability. Differences in performance among the compound classes were noted. Minimum detection limits for the DTD-GC-MS methods were on the order of 0.1 to 1 pg/mm^2 and were as good as or better than those obtained for the solvent extraction method.

Keywords: direct thermal desorption-gas chromatography-mass spectrometry; organic speciation of diesel particulate matter; method comparison

*Corresponding author. Email: lisa.graham@bluesideup.aero
†Present address: University of Canterbury, Department of Chemistry, Private Bag 4800, Christchurch, New Zealand.
‡Present address: Department of Chemistry, Acadia University, Wolfville, NS, Canada B4P 2R6.

ISSN 0306–7319 print/ISSN 1029–0397 online

DOI: 10.1080/03067310903108352
http://www.informaworld.com

1. Introduction

It is well recognised that elevated particulate matter (PM) concentrations in ambient air produce adverse health effects in humans [1]. Identification of the various PM sources and assessment of the chemical composition of the PM from these sources are important steps in the management of air quality.

Much of the literature on the characterisation of the organic carbon (OC) content of PM has been produced using solvent extraction methods [2–5]. The usual procedure for characterisation of the OC content of PM begins with the collection of PM on filter media with or without a size selective inlet (e.g. cyclone or impactor). Often an adsorbent is placed downstream of the filter media to collect semivolatile OC material that either desorbs from the PM collected on the filter during sampling or is present in the gas phase along with the PM. These semivolatile organic compounds are distributed between the gas phase and the particle phase so the filter plus adsorbent is often considered a 'total' sample. The sample is spiked with recovery standards and then extracted using a suitable solvent. The extract is often subjected to silica gel column chromatography to separate it into non-polar, mid-polar and polar fractions to simplify the analysis by conventional liquid injection gas chromatography-mass spectrometry (GC-MS).

Direct thermal desorption-gas chromatography-mass spectrometry (DTD-GC-MS) is a technique that offers the ability to quickly analyse filter based PM samples without requiring the time-consuming tasks of solvent extraction and clean-up [6–9]. The organic compounds collected on the filter are thermally desorbed, cryofocused, and transferred directly to the GC-MS for analysis. The media chosen for sample collection must be thermally stable, to not decompose at the required desorption temperatures. The filter media of choice for this technique is pre-fired quartz, which is known to have a significant positive OC artifact [10–12] after sampling, meaning it also acts as an adsorbent for semivolatile OC in the vapour phase. The magnitude of this positive artifact depends on the composition of the sample stream and the sample collection conditions. Hence, what is found on the filter is not exactly representative of PM suspended in the air, and careful design of the sample collection strategy is necessary to achieve the objectives of the study.

The DTD-GC-MS technique has very high sensitivity and can therefore be applied to much smaller sample quantities than conventional solvent extraction (a few micrograms instead of hundreds of micrograms of PM). The disadvantages include very complex chromatograms and more frequent instrument maintenance, both of which are the result of all the desorbed material being injected into the GC at once. The technique performs better with nonpolar analytes as compared to polar analytes.

A recent validation study of DTD-GC-MS [13] showed excellent recovery of most compounds of interest in ambient and source PM characterisation with the exception of 5-ring and larger polycyclic aromatic hydrocarbons (PAH) from diesel PM. This study suggests the high elemental carbon content of the PM as a possible reason for low recovery. They also showed that detection limits depend not only on the chemical class of compound but also on molecular weight and volatility. A recent comparison of TD-GC-MS with solvent extraction [14] for ambient air PM samples showed reasonable comparison between their two methods for *n*-alkanes, PAHs and petroleum biomarkers.

The DTD-GC-MS technique has been applied to the characterisation of ambient and source emission PM [7,8,15–17]. Coupling to comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry (GC × GC-TOFMS) has been used

to dramatically improve the resolution of the complex chromatograms and provide higher confidence in both compound identification and quantification [18,19].

For the last decade, the Canadian government has been operating an interdepartmental research programme that characterises source emissions and ambient PM with the goal of determining the contribution of transportation sources to ambient PM levels in Canada. Within this programme, all samples have been historically analysed by one laboratory using the solvent extraction GC-MS method described in this study. The DTD-GC-MS methods described in this study have been independently developed and are being used in place of the solvent extraction method for selected parts of this research programme. The motivation for this study was to compare the results obtained from these three independent DTD-GC-MS methods to the results obtained by the solvent extraction method to understand what effects the substitution of the solvent extraction method by any of the DTD-GC-MS methods may have on the programme's data. This paper reports the results from a two phase comparison. Phase 1 of the study involved analysing a mixture of pure compounds commonly found in PM by DTD-GC-MS. Phase 2 involved determining the same suite of compounds in diesel PM samples collected on quartz filters by both DTD-GC-MS and conventional solvent extraction and clean-up with GC-MS analysis.

2. Experimental

2.1 *Study design*

The three DTD-GC-MS systems were located in laboratories at Natural Resources Canada (NRCan CETC-Ottawa) and Environment Canada (EC-ESTC and EC-AQRD). The solvent extraction method was performed by Environment Canada. The present study was designed to evaluate two aspects of performance of the three DTD-GC-MS systems and the conventional solvent extraction method: (1) identification and (2) quantification of a suite of *n*-alkanes, pristane and phytane, petroleum biomarkers (hopanes and steranes) and PAH commonly found in both ambient air and source emission PM samples. The instrumentation installed in each laboratory is summarised in Table 1. Each laboratory was responsible for its own calibration using materials and methods that have been previously developed and applied. Each laboratory, however, was asked to implement a similar set of instrument parameters, as summarised in Table 1, for the DTD-GC-MS instrumentation to facilitate comparison of the results.

2.1.1 *Phase 1: Analysis of unknown solution*

A liquid solution of the target analytes, with concentrations in the range of 0.1 to $1\,\mu g\,mL^{-1}$, was provided to each laboratory. The exact concentration was unknown to the laboratories. The solution was spiked onto blank filters and analysed. Each laboratory provided its own blank filters for this phase. The solution was spiked at two levels which represented the lower and upper limits of the laboratories' calibration ranges. This phase of the study provided information on the differences in performance of each DTD-GC-MS method compared to the solvent extraction method on a well defined sample that is free of the matrix usually associated with the real filter based PM samples.

Table 1. Comparison of DTD-GC-MS instrumentation in the three laboratories.

	EC-ESTC	EC-AQRD	NRCan CETC-Ottawa
Thermal desorption-gas chromatograph-mass spectrometer calibration method	Gerstel TDU-CIS4 Agilent 6890 Agilent 5975 Internal standard 6 points 0.02–1 $\mu g\,L^{-1}$ all compounds Linear fit	Gerstel TDS-CIS4 Agilent 6890N Agilent 5973 Internal standard 5 points Biomarkers: 0.01–22.8 $\mu g\,L^{-1}$ PAH: 0.01–5.01 $\mu g\,L^{-1}$ *n*-Alkanes: 0.04–10.18 $\mu g\,L^{-1}$ Linear fit	Gerstel TDU-CIS4 Agilent 6890 Agilent 5975 External standard 5 points 0.125–2 $\mu g\,L^{-1}$ (low level spike) 1–10 $\mu g\,L^{-1}$ (high level spike) 0.125–60 $\mu g\,L^{-1}$ (diesel PM filter) (8 point) Quadratic fit
Sample size	One 8.1 mm diameter punch	One 7 mm diameter punch	Two 7 mm diameter punches
Surrogates (Recovery standards) and Internal standards	**Internal standards** Spike volume = 1 µL *n-Alkanes* (1 $\mu g\,L^{-1}$) *n*-Tetracosane-D50 *PAHs* (1 $\mu g\,L^{-1}$) Naphthalene-D8 Acenaphthene-D10 Phenanthrene-D10 Chrysene-D12	**Internal standards** Spike volume = 2 µL *n-Alkanes* (5 $\mu g\,L^{-1}$) *n*-nonadecane-D40 *n*-tetracosane-D50 *n*-triacontane-D62 *PAHs* (1 $\mu g\,L^{-1}$) Phenanthrene-D10 Chrysene-D12	**Surrogates** Spike volume = 2 µL *n-Alkanes* (0.25 $\mu g\,L^{-1}$) *n*-Dodecane-D26 *n*-Hexadecane-D34 *n*-Eicosane-D42 *n*-Tetracosane-D50 *n*-Triacontane-D62 *n*-Dotriacontane-D66 *n*-Hexatriacontane-D74

	Biomarkers ($2\,\mu g\,L^{-1}$) 5α(H)-Androstane	*Biomarkers* ($1\,\mu g\,L^{-1}$) Phenyldodecane	*PAHs* ($0.25\,\mu g\,L^{-1}$) Naphtalene-D8 Phenanthrene-D10 Pyrene-D12 Benz(a)anthracene-D12 Perylene-D12 *Biomarkers* ($0.25\,\mu g\,L^{-1}$) $\beta\beta$-Hopane
DTD-GC-MS Instrument parameters used by all three laboratories	**Gerstel TDU and TDS** Initial Ramp Final **Cold injection system (CIS)** Initial T Ramp Final T **GC-MS** Initial Ramp Final T Column MSD	 T20°C 25°C/min T320°C −60°C 720°C/min 325°C T60°C hold for 3 min 5°C/min 320°C hold for 10 min DB-5MS 0.25 mm × 30 m × 0.25 µm film flow $1.2\,mL\,min^{-1}$ Electron impact Selected ion monitoring mode (SIM)	

For comparison, the liquid sample was analysed directly by GC-MS and also spiked onto filters which were then extracted, cleaned up and analysed in the same way as the diesel PM samples.

2.1.2 *Phase 2: Analysis of diesel PM samples*

A late model heavy-duty diesel engine was used as the source of emissions. Samples of diesel PM were collected on pre-fired quartz filters using standard dilution sampling techniques. Multiple simultaneous samples were collected over several repeat engine tests. Each filter was divided into quarters. Each laboratory was provided with one quarter of each of three different filters as well as a blank filter prepared in the same manner as the filters used for sample collection. Each laboratory analysed the samples following their own procedures but with the common set of instrument parameters. The DTD-GC-MS results were compared to the solvent extraction results, thus providing information on the ability of the laboratories to correctly identify and quantify the target analytes in real samples.

2.2 *Liquid sample preparation*

A liquid sample was prepared in iso-octane using *n*-alkanes, pristane and phytane, petroleum biomarkers (Chiron AS, Norway) and a solution of 16 priority PAHs (Supelco, USA). This sample included the target analytes at concentrations in the range of 0.1 to $1\,\mu g\,mL^{-1}$. Recovery surrogates and/or internal standards, if needed for quantification, were added by the laboratories conducting the analyses.

2.3 *PM filter sample collection*

Diesel PM samples were collected from a 2004 Cummins ISM280 diesel engine operating in its certified configuration which included an oxidation catalyst. The fuel used was a commercial ultra-low sulphur diesel (<15 ppm sulphur). The entire volume of exhaust produced by the engine was collected and diluted using a constant volume sampling system following standard procedures [20]. The engine was operated at a constant speed and load.

Four simultaneous samples of the dilute exhaust were collected on filters with URG-2000-30ENB cyclones. Each cyclone was fitted with a 90 mm diameter 2-stage filter pack. Two of the filter packs contained only a primary quartz filter (Pall Tissuquartz™ Filters, 2500 QAT-UP) to collect PM. The other two filter packs contained a primary Teflon membrane filter (Pall Zefluor™ Membrane, 2 μm pore size) with a quartz filter downstream of the Teflon filter. The primary Teflon membrane filter was used to determine the dilute exhaust PM concentration by gravimetry. The quartz filters were not subjected to gravimetric analysis as they tend to lose fibres in the filter holders during sample collection and are difficult to get stable humidity equilibration and still retain low OC blank levels. Typically, the secondary quartz filter collected downstream of the Teflon membrane filter is used to correct for the organic carbon adsorption artifact.

Samples were collected at a flow rate of 60 lpm for 20 minutes as this was the maximum stable flow rate that could be obtained through the Zefluor membrane filters with the equipment used. Since the cyclones that were used achieve a $PM_{2.5}$ cut at 91 lpm, the actual size cut achieved at 60 lpm was 3.5 μm. Eight sets of filters were collected. Four of the

primary quartz filters were cut into quarters and each laboratory received one quarter of each of three separate filters. The remaining filters were archived in a freezer at −20°C for future use.

A fifth filter sample was also collected using the standard secondary dilution sampling technique specified by emissions certification test procedures [20]. This sample was collected on a Teflon bound borosilicate glass material (Pall EMFAB^{TM} TX40HI20WW) and was subjected to gravimetric analysis. The dilute exhaust concentrations determined by the three filters were within 10% of each other as shown in Table A in the Supplementary Material (available online). We have assumed that the quartz filters, for which there is no gravimetric data, are similarly loaded. Since each laboratory received a portion of each of the same filters, whether each filter is identical or not to the others will not affect the results provided the filters have homogeneous PM deposits.

2.4 *Solvent extraction GC-MS analysis*

2.4.1 *Solvent extraction*

Three blank filters were spiked with 50 µL and three blank filters were spiked with 250 µL of the unknown solution. One-fourth of a 90 mm diameter filter comprised each diesel PM sample. Each filter was crumbled into a 30 mL pressure vessel and spiked with recovery surrogates (d-alkanes, $\beta\beta$-hopane, and d-PAHs) and allowed to dry for 30 minutes. Pre-cleaned sand was added to the vessel to fill the void volume. The samples were loaded into a pressurised solvent extraction system (DIONEX ASE 300) with the first sample being an instrument blank. This ensured that the system was warmed up to operating temperature for the sample extractions thus maximising recoveries. The ASE parameters were as follows: solvent: dichloromethane; cell pressure and temperature: 1500 psi at 100°C; extraction time: 5 min heat up and 5 min static; cycles: 3; flush 120%; nitrogen purge at 100 psi for 100 s.

The sample extracts were removed from the ASE and the filters were extracted again using the same parameters replacing the dichloromethane with methanol. This extract was collected separately and archived. Control samples consisting of the pressure vessel filled with sand, were spiked with the recovery surrogates as above and native *n*-alkanes, petroleum biomarkers and PAHs were also added. The controls were processed in the same manner as the samples.

2.4.2 *Sample clean-up*

Dichloromethane extracts were concentrated to 5 mL and exchanged to hexane with a final volume of 5 mL. A deactivated silica column containing 5% water (w/w) was used for clean-up. The column was pre-washed with 10 mL of hexane. A 125 mL flask labelled F1 was placed under the column, the sample was transferred to the column and it was allowed to go to bed dry. Three 1 mL hexane rinses of the original sample flask were added to the column. The column was then eluted with 5 mL hexane and allowed to go just to bed dry. Then 5 mL of 50/50 hexane/acetone was added to the column and allowed to go to bed dry. The F1 flask was changed to a centrifuge tube labelled F2. Ten mL of 50/50 hexane/acetone was added to the original sample flask and swirled to rinse. This volume was then added to the column and allowed to go to bed dry. The F2 tube was changed to a

20 mL vial labelled F3. Then 20 mL of methanol was added to the original flask, swirled and transferred to the column.

Fraction F1 (non-polars – alkanes and biomarkers) was gently concentrated to 3–5 mL by rotary evaporation and transferred with 3 hexane rinses to a centrifuge tube pre-calibrated to 1.0 mL with iso-octane. The extract was concentrated to approximately 1 mL using a gentle purge of nitrogen and the internal standards (phenyldodecane and 5α-androstane) were added. Fraction F2 (mid-polars – PAHs) was concentrated to approximately 0.45 mL using a gentle purge of nitrogen and 50 µL of the internal standard (d10-fluoranthene) was added. Fraction F3 (polars) and the ASE methanol extract were archived. Only Fractions F1 and F2 were analysed for this study.

2.4.3 *GC-MS analysis*

The instrument parameters for the analysis of the non-polar (F1) and mid-polar (F2) fractions are summarised in Table 2.

The *n*-alkanes and biomarkers were calibrated from 10 to 1000 µg L^{-1} with reference to their internal standards (phenyldodecane and 5α-androstane) and a quadratic fit was used to establish the calibration curves. A 1 µL volume of the extract was injected in an Agilent 6890GC/5973MS. An instrument blank was analysed right after the calibration run and the GC-MS was proven to be contamination free. A check standard sample was also inserted into the analysis sequence to periodically verify the validity of the calibration curve.

The PAH calibration was done at seven concentration levels from 0.01 to 16 ng/µL with reference to the internal standard d10-fluoranthene. Linearity was confirmed with a correlation coefficient better than 0.995. A volume of 1 µL was injected on-column into the Agilent 6890GC/5973MS. Check standards were run after every seven samples and at the beginning and end of each day to verify the calibration curve.

The internal standard quantification was used to correct for variables such as differences in final extract volume and instrument drift. Results were also corrected for solvent extraction efficiency using the measured recovery of surrogates added prior

Table 2. GC-MS analysis parameters for the solvent extraction method.

	Non-polar *n*-alkanes and biomarkers	Mid-polar PAHs
GC	Agilent 6890	Agilent 6890
Inlet	300°C Splitless	275°C Split
Column	DB-5MS	DB-XLB
	0.25 mm × 30 m × 0.25 µm film	0.25 mm × 30 m × 0.25 µm film
	Flow: 1.5 mL min^{-1} Helium	Flow: 2 mL min^{-1} Helium
Temperature	Initial T: 50°C	Initial T: 90°C hold 5 min
	Ramp: 8°C min^{-1}	Ramp: 20°C min^{-1} to 200°C
	Final T: 320°C hold 15 min	2.5°C min^{-1} to 250°C
		1.5°C min^{-1} to 283°C
		Final T: 283°C hold 6 min
MSD	Agilent 5973, EI	Agilent 5973, EI
	Selected ion monitoring (SIM)	Selected ion monitoring (SIM)

to extraction. Compounds with surrogate recoveries less than 30% were reported as NQ (not quantifiable).

2.5 *Thermal desorption analysis*

2.5.1 *Blank filters and analysis of liquid samples*

Each laboratory prepared its own blank filters for instrument calibration and analysis of the unknown liquid sample. The pertinent details are summarised below.

EC-ESTC: Quartz fibre filters (PALL Inc. Mississauga, ON) were pre-fired at 900°C for a minimum of 5 h to remove any residual carbon contamination. The pre-fired filters were wrapped in cleaned aluminium foil (baked in an oven at 150°C for 5 hr), put in sealed polyethylene bags and stored in a freezer (−15°C) until ready for use. A single punch (8.1 mm diameter) cut from the filter was transferred to a thermal desorption (TD) tube. The calibration solutions or unknown liquid sample, along with the internal standard were spiked on the punch for analysis.

EC-AQRD: 7 mm diameter punches of quartz fibre filters (PALL) were pre-fired at 800°C for 2 h in a furnace. When the furnace had cooled to 100°C, they were removed and put into a pre-cleaned glass bottle for storage in a freezer prior to use. The calibration solutions or unknown liquid sample were spiked on the pre-fired filter punch and put into the TD tube for analysis. The TD tubes were cleaned using laboratory detergent and pure water, then conditioned in a controller (Gerstel Aux-Controller 163) at 350°C with Helium flow for 2 h.

NRCan CETC-Ottawa: AH934 glass fibre filters were pre-cleaned using an ASE apparatus and stored in a muffle furnace at 250°C until ready for use. Punches (7 mm diameter) cut from the filter were transferred to TD fritted tubes. TD tubes with the punches were conditioned using a Gerstel TC2 conditioning unit for 4 h (100 mL min^{-1} N_2, 350°C) then transferred into a muffle furnace (250°C) until used. Tubes were allowed to cool to room temperature, after which the unknown solution was spiked with a syringe on the pre-cleaned punch (inside the tube) for analysis. TD tubes that were used for PM filter samples were cleaned using the same procedure, except without a blank filter punch.

2.5.2 *Calibration*

Each laboratory followed their calibration procedures that were independently developed and implemented. These procedures are briefly summarised below and specific details are summarised in Table 1.

EC-ESTC: Standard solutions were prepared in iso-octane using reference materials including *n*-alkanes, pristane and phytane, petroleum biomarker compounds (hopanes and steranes) and PAHs (Chiron AS, Norway). The calibration solution and internal standard solution were spiked on the filter punch (inside the desorption tube) separately using a syringe and analysed.

EC-AQRD: Standard solutions were prepared from NIST SRMs 2260a (PAH) and 2266 (petroleum biomarkers) and *n*-alkanes (Sigma Aldrich, USA). A known amount of standard solution was spiked on the pre-cleaned filter punches (inside the tube) using a syringe and analysed. The calibration solution and internal standard solution were spiked on the filter punch separately.

NRCan CETC-Ottawa: Standard solutions of *n*-alkanes, pristane, phytane, petroleum biomarkers (all from Chiron AS, Norway) and PAHs (UltraScientific, USA) as well as selected deuterated compounds were prepared in hexane. A known amount of standard solution (0.25 μg mL^{-1} mix of deuterated PAHs, deuterated *n*-alkanes and *ββ*-hopane) was spiked on the pre-cleaned filter punches (inside the tube) using a syringe and then analysed. An external calibration was prepared for the native and deuterated compounds. Calibration ranges were different depending on the sample being analysed. The liquid unknown solution was analysed directly without the use of the deuterated standard.

2.5.3 *Diesel PM filter analysis*

The three diesel PM filter samples were analysed by each laboratory using their standard procedures that are briefly summarised below.

EC-ESTC: A single punch of the diesel PM filter was spiked with the internal standard solution and analysed. Recovery surrogates were not used. Each filter was analysed in triplicate.

EC-AQRD: A single punch of the diesel PM filter was spiked with the internal standard solution and analysed. Recovery surrogates were not used. Each filter was analysed in triplicate.

NRCan CETC-Ottawa: Two punches of the diesel PM filter were spiked with a deuterated compound standard (2 μL of a 0.25 μg mL^{-1} mix of deuterated PAHs, deuterated *n*-alkanes and *ββ*-hopane solution) and results were corrected for recovery of these compounds. An internal standard was not used as the method has an external calibration. Each filter was analysed in triplicate.

3. Results and discussion

3.1 *Analysis of unknown liquid sample*

For the comparison of unknown liquid sample results, the accuracy of a method for a given target analyte is defined as the difference between the mean of the three repeat analyses and the target concentration of the unknown liquid sample. The relative accuracy of a method for a given target analyte is defined as the accuracy divided by the target analyte concentration in the unknown liquid sample. The relative accuracy for a given target analyte can also be expressed as a percentage.

As one of our objectives was to compare the results of the DTD-GC-MS methods with those of the solvent extraction method, a second measure of accuracy for a given target compound can be defined for the DTD-GC-MS methods using the results of the solvent extraction method for the unknown solution as the reference instead of the target concentration. A corresponding relative accuracy is defined as the accuracy divided by the solvent extraction result. This accuracy can also be expressed as a percentage.

The precision of a method for a given target analyte is defined as the standard deviation of the mean of three repeat analyses. The relative precision for a given target analyte is defined as the standard deviation divided by the mean and can also be expressed as a percentage.

Since each method determines up to 61 individual analytes, the following two metrics are defined in an effort to summarise the performance of an individual method and to simplify comparison between methods. The average relative accuracy of a method

(*vs*. the target concentration or *vs*. the solvent extraction method) is defined as the average of the relative accuracies of all target analytes of a given class or for all target analytes determined by the method, *without* respect to sign (i.e. the average of the absolute value of the target compound accuracies). The smaller the value of the average relative accuracy, the closer the method is to achieving the results obtained by the reference to which it is compared. The range of this average relative accuracy is defined by the minimum and maximum relative accuracies. The average relative precision of a method is defined as the average of relative precisions of all target analytes of a given class or for all target analytes determined by the method. The range of this average relative precision is defined by the minimum and maximum relative precisions.

The target analytes and their concentrations in the unknown liquid sample are shown in Table B of the Supplementary Material. Also shown in Table B are the average concentration, relative accuracy (RA%) and relative precision (RP%) for each target analyte obtained from three replicates of the direct analysis of the liquid sample. The average relative accuracy and relative precision metrics are summarised in Table 3. The average relative accuracies were similar for the *n*-alkanes and PAHs (7% and 8%, respectively) but higher for the biomarkers (17%). The average relative precision was better for the *n*-alkanes and biomarkers (1%) than for the PAHs (5%). The *n*-alkane and PAH results are within the usual performance limits of liquid sample injection

Table 3. Average relative accuracies and average relative precisions for the analysis of the unknown liquid sample. Average (min–max).

	Average relative accuracy (%)[†]		
	Reference target concentration	Reference solvent extraction	Average relative precision (%)
Alkanes			
Direct liquid analysis	8 (0–26)		1 (0–3)
Solvent extraction*	60 (0–184)		63 (7–149)
NRCan CETC-Ottawa	5 (0–13)	36 (1–67)	9 (2–18)
EC-ESTC	6 (0–49)	36 (0–74)	9 (2–27)
EC-AQRD	19 (6–29)	51 (11–73)	6 (2–25)
Biomarkers			
Direct liquid analysis	17 (1–35)		1 (1–2)
Solvent extraction	20 (6–30)		5 (2–10)
NRCan CETC-Ottawa	5 (0–11)	22 (4–37)	4 (2–5)
EC-ESTC	16 (12–20)	9 (1–24)	4 (1–9)
EC-AQRD	32 (4–69)	28 (2–110)	7 (3–12)
PAH			
Direct liquid analysis	7 (2–23)		5 (0–16)
Solvent extraction	15 (1–36)		6 (2–12)
NRCan CETC-Ottawa	9 (1–44)	35 (6–124)	7 (3–13)
EC-ESTC	5 (0–13)	15 (0–45)	7 (2–18)
EC-AQRD	31 (1–150)	42 (1–178)	16 (2–77)

Notes: [†]Average relative accuracy is calculated *without* regard for sign, i.e. using absolute values of target compound accuracies.
*Contamination has affected alkane results from the solvent extraction method.

analyses of this type. The biomarker compound difference from target is larger and may influence the interpretation of method comparison results.

3.1.1 *Solvent extraction method*

For each target analyte, the average of 3 replicate analyses of the low and high level spiked filters, the relative accuracies and relative precisions are summarised in Table B in the Supplementary Material. The very high *n*-alkane results prompted a search for sources of contamination that will be discussed later. There was no apparent contamination affecting the PAH and biomarker compounds. Naphthalene and acenaphthylene had surrogate recoveries less than 30% and are reported as NQ. For the biomarkers and PAH, the relative accuracy for each target compound was similar to the relative accuracy obtained for the direct analysis of the unknown liquid sample (Figure A in Supplementary Material). This was expected since the same analytical methods were used.

The *t*-test (two tailed, assumed equal variances, $\alpha = 0.05$) was used to determine if there was a significant difference between the results for the low level and high level spike. The results are summarised in Table 4 and illustrated in Figure B in the Supplementary Material. For more than half the reported compounds, there was no statistically significant difference between the low level and high level spiked filters. The individual results for the low and high level spiked filters were averaged for comparison to the DTD-GC-MS results. This comparison is shown in Figure 1 and the abbreviations used for analyte names in this figure are given in Table B of the Supplementary Material. The average relative accuracies (target concentration as reference) and the average relative precisions for the three compound classes for the solvent extraction method are summarised in Table 3.

3.1.2 *DTD-GC-MS methods*

Two of the laboratories (NRCan CETC-Ottawa and EC-AQRD) used similar spike volumes of 2 or 2.5 μL for low level and 10 μL for high level while the third laboratory

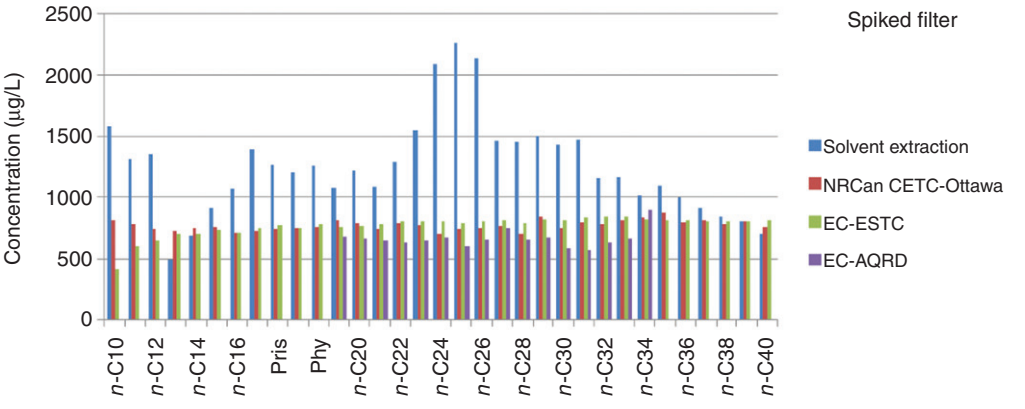
Table 4. Significance test results comparing low level and high level spiked filters and comparing the three DPM filter samples within a method.

There is no statistically significant difference (*t*-test, assumed equal variances, $\alpha = 0.05$) between low and high level spikes for *x* of *y* target analytes reported.
The total number of target analytes considered by the methods is *n*.

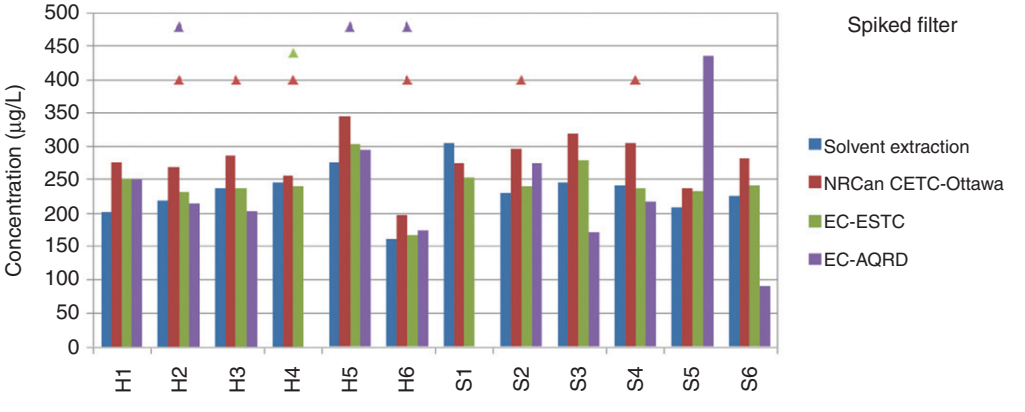
	Solvent extraction	NRCan CETC-Ottawa	EC-ESTC	EC-AQRD
Alkanes ($n = 33$)	Not tested	20 of 31	13 of 33	0 of 16
Biomarkers ($n = 12$)	8 of 12	12 of 12	4 of 12	1 of 10
PAH ($n = 16$)	7 of 14	8 of 15	10 of 16	4 of 16

There is no statistically significant difference (ANOVA, $\alpha = 0.05$) among the three DPM filter samples for *x* of *y* target analytes reported.
The total number of target analytes considered by the methods is *n*.

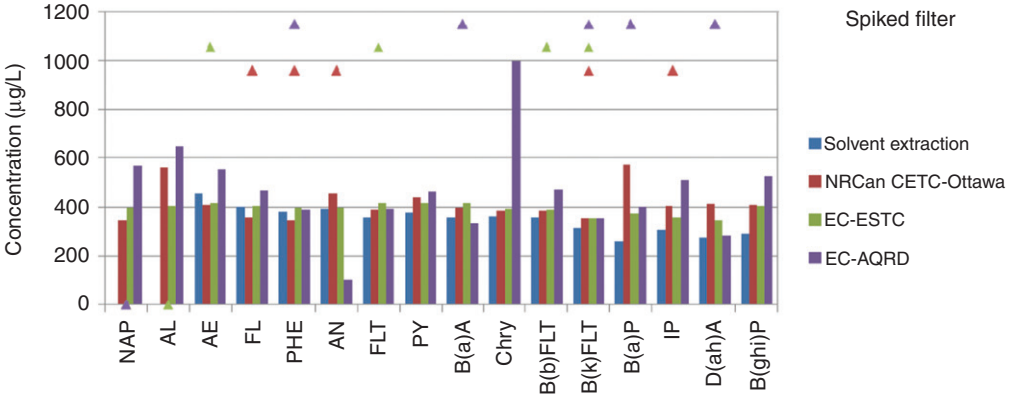
	Solvent extraction	NRCan CETC-Ottawa	EC-ESTC	EC-AQRD
Alkanes ($n = 33$)	Not tested	16 of 27	12 of 28	12 of 16
Biomarkers ($n = 12$)	Not tested	7 of 12	12 of 12	4 of 8
PAH ($n = 16$)	Not tested	7 of 16	12 of 16	9 of 12



(a) Alkanes



(b) Petroleum biomarkers



(c) PAH

Figure 1. Comparison of spiked filter results for the solvent extraction and the three DTD-GC-MS methods. Triangles indicate analytes[†] for which the DTD-GC-MS method is not statistically significantly different from the solvent extraction result.
Note: [†]Abbreviations for analyte names are defined in Table B.

(EC-ESTC) used lower spike volumes of 0.5 and 2.5 μL. Two of the laboratories (NRCan CETC-Ottawa and EC-ESTC) used calibration standards for *n*-alkanes and biomarkers sourced from the same supplier (Chiron) with the same lot numbers but different PAH standards (Ultra Scientific and Chiron). The third laboratory (EC-AQRD) calibrated with NIST Standard Reference Materials (SRM) 2260a for PAHs and 2266 for biomarkers and prepared an *n*-alkane standard from materials purchased from Sigma-Aldrich. This laboratory also reported a subset of the target analytes. The averages of three replicate analyses of the unknown liquid sample as reported by the three laboratories are summarised in Table C and compared in Figure B in the Supplementary Material. The relative accuracies for each target analyte for the three methods are compared in Figure C in the Supplementary Material.

The *t*-test (two tailed, assumed equal variances, $\alpha = 0.05$) was used to determine if there was a statistically significant difference between the low and high level spiked filters. The results are summarised in Table 4 and illustrated in Figure B in the Supplementary Material. The results indicate that for some laboratories and analytes the difference between low and high level spikes is statistically significant. While the differences may be statistically significant, the differences are not large in the context of this type of analysis. The differences are generally in the range of 5–15% and are found to be statistically significant because of the very good precision of repeat analyses. This is particularly evident for EC-AQRD where precisions were on the order of 2% and average differences between low and high level spike were on the order of 15%. The individual low and high level spiked filter results were averaged for comparison to the solvent extraction method. This comparison is shown in Figure 1. The average relative accuracies (using both target concentration and the solvent extraction results as reference) and the average relative precisions for the three DTD-GC-MS methods are summarised in Table 3.

Two of the three laboratories (NRCan CETC-Ottawa and EC-ESTC) achieved similar average relative accuracies compared to the target concentration ($\leq$16%) while the third laboratory (EC-AQRD) consistently achieved poorer relative accuracies (19–32%). The average relative precisions were similar among the laboratories ($<$10%, with the exception of EC-AQRD PAH) and were similar to the solvent extraction method. The average relative precisions were similar across compound class, with the exception of a few compounds (C_{10} to C_{13} *n*-alkanes when reported, the last 4 PAHs for EC-AQRD). Each of the laboratories showed slightly different performances by class of compound (Table 3). Only one laboratory reported the most volatile *n*-alkanes (C_{10} to C_{13}) for both spike levels, but with large negative differences from target, and with precision that degraded as the amount of material decreased. All three laboratories reported naphthalene in the high level spike, but only two of three reported it in the low level spike. The reason for not reporting the low level spike was a co-elution with the deuterated dodecane surrogate and its interference with the quantification ion (m/z 128) used for naphthalene. These two peaks were sufficiently resolved in the high level spike to allow for quantification.

The differences in calibrations are evident in the results for spiked filter samples. The NIST standards used by EC-AQRD have different concentrations for each component and as a result, much larger calibration ranges had to be prepared to quantify the PAHs and biomarkers. The variability of the replicate analyses on the EC-AQRD system are small ($\sim$2%) but the large difference from the target value for some target analytes suggests that the instrument response may not be linear over the calibration range used. The concentration difference between lowest and highest compounds in the EC-AQRD calibration standard solution is a factor of 37 for the biomarkers and a factor

of 7 for the PAHs. The calibration range for the EC-AQRD method may also be partly responsible for the large number of analytes that showed statistically significant differences between the low level and high level spiked filters.

3.2 *Analysis of diesel PM filters*

3.2.1 *Solvent extraction method*

One quarter of each of 3 different diesel PM filter samples was processed using the solvent extraction and clean-up procedures described above. The average of these three analyses and the relative precision for each target analyte are summarised in Table D of the Supplementary Material. The contamination that invalidated the unknown liquid sample spiked filter results was also present in the diesel PM filter results, but appeared to be sufficiently separated from the *n*-alkane distribution of the diesel PM that the blank corrected results obtained were deemed useful. These results are shown in Figure 2 and the abbreviations used for analyte names in this figure are given in Table B of the Supplementary Material.

The average relative precisions by compound class for this analysis are summarised in Table 5. The C_{10} to C_{13} *n*-alkanes were not quantified because the surrogate recoveries were less than 30%. The results for the C_{15} to C_{21} *n*-alkanes were quite repeatable with individual compound precisions ranging from 0.9% to 12%, with an average of 7%. The results for the C_{22} to C_{24} *n*-alkanes were highly variable and the C_{25} to C_{40} *n*-alkanes were not quantified due to the contamination. Only 5 of the 16 target PAHs were detected in the sample. Naphthalene was detected but not quantified because surrogate recovery was less than 30%. The precision for these 4 compounds ranged from 5% to 20% with an average precision of 11%. All but one of the 12 target biomarkers were detected and the variability was similar to that obtained for the PAHs. Note that the PAHs and biomarkers are present in concentrations substantially lower than the *n*-alkanes yet the methods achieve similar analytical performance. This difference in performance is likely due to the lack of specificity with which the *n*-alkanes can be identified in the presence of other saturated and unsaturated hydrocarbons which comprise the unresolved complex material (UCM) that is characteristic of these samples and the resulting difficulties in integrating peaks which are superimposed on this UCM.

3.2.2 *DTD-GC-MS method*

Each of the three quarter filters was analysed in triplicate by each of the laboratories. The average concentration (pg/mm^2) and the relative precision for each target analyte for the three filters are summarised in Table D of the Supplementary Material. Analysis of Variance (ANOVA, $\alpha = 0.05$) was used to determine if the results were similar among the three filters within each method for each target analyte. The results are summarised in Table 4 and illustrated in Figure D in the Supplementary Material. These results show that the three filters were reasonably similar in composition, allowing the results to be averaged. The average results are compared to the solvent extraction results in Figure 2.

Two of the laboratories (NRCan CETC-Ottawa and EC-ESTC) had similar average precisions and precisions by compound class which, in some cases, were notably higher than the third laboratory (EC-AQRD). There are two reasons for this difference which can be traced to the *n*-alkanes. A narrower range of *n*-alkanes was reported by EC-AQRD

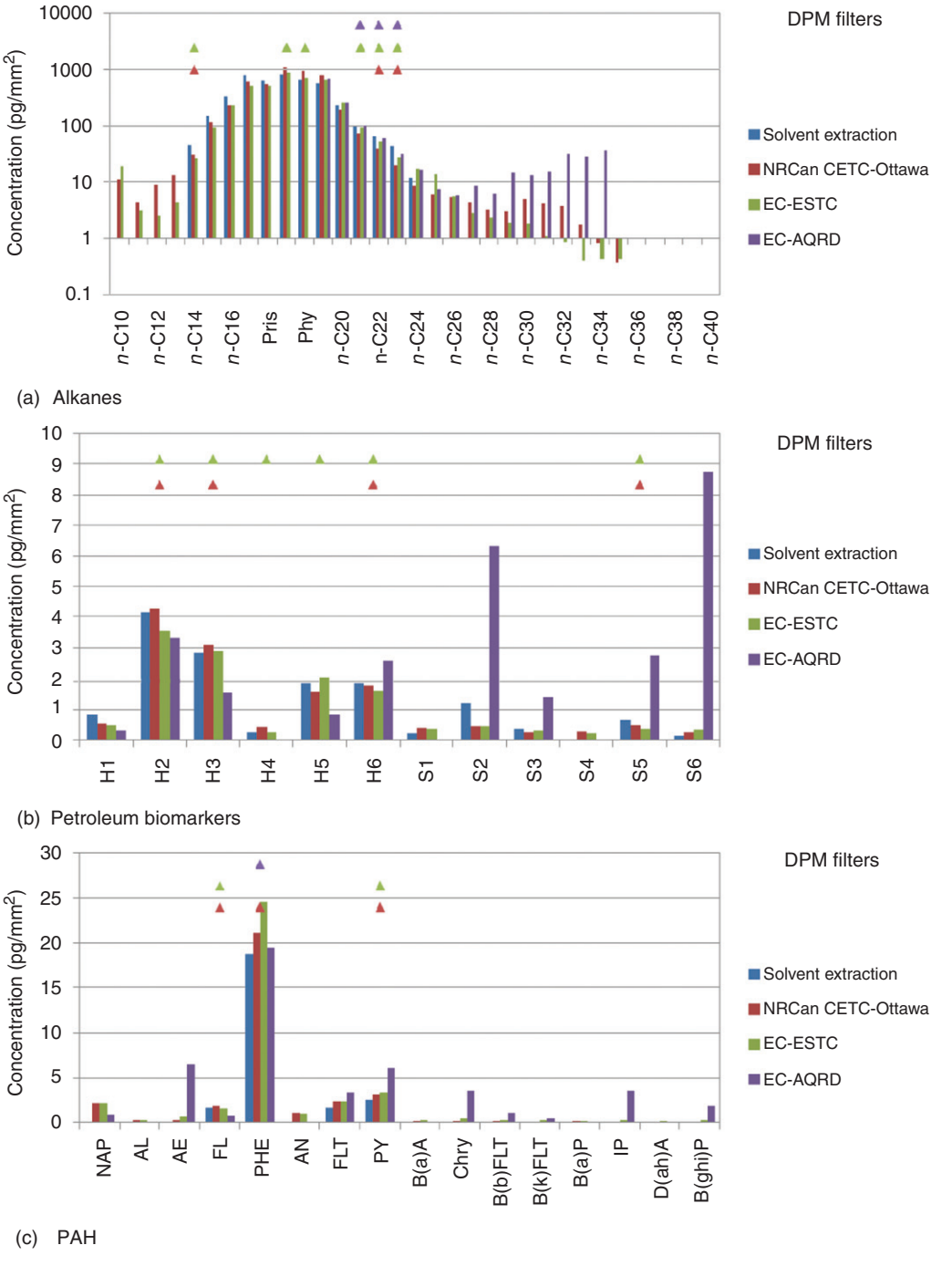


Figure 2. Comparison of diesel PM filter results for the solvent extraction and the three DTD-GC-MS methods. Triangles indicate analytes† for which the DTD-GC-MS method is not statistically significantly different from the solvent extraction result.
Note: †Abbreviations for analyte names are defined in Table B.

Table 5. Average relative accuracy (solvent extraction as reference) and average relative precision achieved for the analysis of diesel PM filters by solvent extraction and thermal desorption. Results are reported as average (min-max).

	Average relative accuracy† (%)					
	Reference solvent extraction			Average relative precision (%)		
	Alkanes	Biomarkers	PAH	Alkanes	Biomarkers	PAH
Solvent extraction				18 (1–98)	12 (6–22)	13 (5–22)
NRCan CETC-Ottawa	31 (14–55)	40 (2–99)	39 (13–88)	18 (0–72)	10 (0–31)	21 (1–85)
EC-ESTC	24 (4–44)	39 (1–168)	42 (8–88)	18 (1–72)	9 (1–32)	16 (1–92)
EC-AQRD	19 (3–43)	907 (20–6889)	82 (4–147)	7 (1–46)	8 (0–54)	11 (0–55)

Note: †Average relative accuracy is calculated *without* regard for sign, i.e. using absolute values of target compound accuracies.

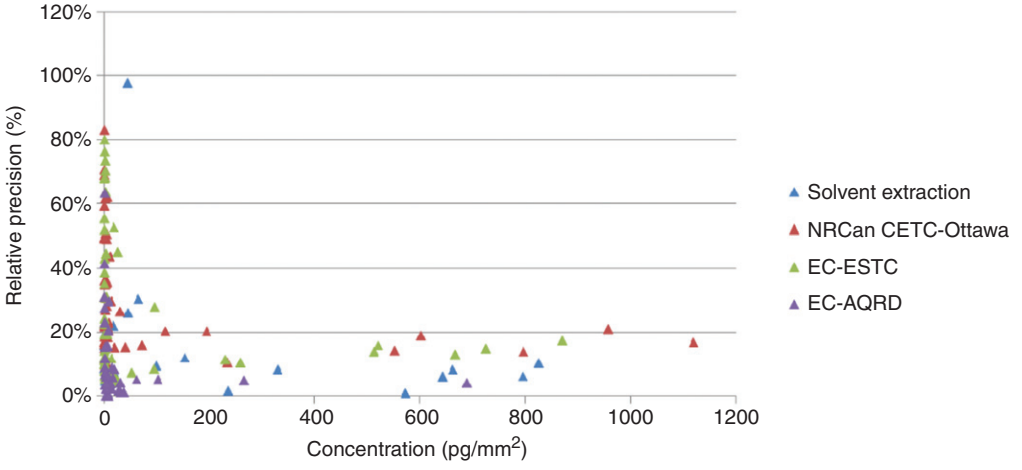


Figure 3. Analytical precision as a function of analyte concentration for diesel PM samples.

(C_{19} to C_{34}). This laboratory chose to eliminate from their method those compounds for which the results are most variable due to volatility and low concentration. This decision also resulted in omission of the majority of the *n*-alkanes emitted by diesel engines from the *n*-alkane source profile. The concentrations for the C_{29} to C_{34} *n*-alkanes reported by EC-AQRD are between 3 and 10 times higher than those reported by the other two laboratories and have correspondingly better precisions. The results of the two EC labs were similar for PAHs and the petroleum biomarkers. The poorer precisions seen for NRCan CETC-Ottawa are likely due to their choice of external calibration with recovery correction rather than a true internal calibration.

As shown in Figure 3, for the majority of components with concentrations above 25 pg/mm^2, the relative precision is less than 25%. Larger variability is observed for components with concentrations below 25 pg/mm^2. There are many factors such as loss

due to volatility, difficulty in reliably identifying and integrating small peaks on a high baseline, and background contamination that can result in poorer precision. With TD methods, since all components are transferred onto the GC, there is more material present to cause interferences and co-elutions with target compounds.

3.3 *Alkane contamination study*

The solvent extraction results for the *n*-alkanes suggested contamination had occurred during storage and/or sample handling that resulted in a significant increase in alkane content of the samples. There was no apparent effect on the PAH or biomarker content. The details of the experiments conducted in an effort to identify the source of the contamination are summarised in the Supplementary Material. The contamination was ultimately attributed to the cleaning of the small filters inserted in the ASE cells.

A set of filters spiked with the unknown liquid sample was also analysed by a different solvent extraction method implemented by NRCan CETC-Ottawa and showed good agreement with the target concentration (Figure G in the Supplementary Material).

As part of the search for the source of contamination, a second set of spiked filter and diesel PM samples were also analysed for PAHs and biomarkers by the solvent extraction method. These results compare well to the first solvent extraction samples as shown in Figure H in the Supplementary Material, indicating that the detection limits were sufficient for the single diesel PM filter used in the first set of samples.

3.4 *Comparison of DTD-GC-MS methods*

ANOVA ($\alpha = 0.05$) was used to determine if there was a statistically significant difference among the three DTD-GC-MS methods for each of the analytes reported by all 3 methods. This test does not identify which of the methods is (or are) responsible for the difference, it only tests to see if the results are comparable among the methods. This test also does not compare the DTD-GC-MS methods to the reference method. The results obtained from the ANOVA analysis are as follows. From the spiked filter analyses, for the 16 *n*-alkanes reported by all three methods, 14 of them showed a statistically significant difference among the three methods. For the biomarkers, all 10 of the compounds reported by all methods showed a statistically significant difference. For the PAHs, 12 of the 14 analytes reported by all three methods showed a statistically significant difference. From the diesel PM filter analyses, 15 of 16 *n*-alkanes, all 10 of the biomarkers and all 12 of the PAHs reported by all three methods showed a statistically significant difference. The methods are clearly not interchangeable as defined by this criterion; however, this finding is not as discouraging as it may sound.

The comparison that is most important to this study is the comparison of each DTD-GC-MS method to the solvent extraction method. This distinction is subtle but important. For example, if one of the DTD-GC-MS methods reports consistently 10% high compared to the solvent extraction method, another consistently 10% low compared to the solvent extraction method, then the difference among the DTD-GC-MS methods as tested by the ANOVA may be statistically significant, but the difference from the reference solvent extraction method may not be statistically significant. Because the methods have good precision, small differences in accuracy can be detected by the ANOVA procedure. Equally important is the possibility that while two methods may be statistically

significantly different, from an operational perspective, the magnitude of the difference may be tolerable. This situation is likely quite often encountered, especially with current instrumental and automation improvements in analytical techniques reducing the variability among repeat analyses. The importance of accurate calibration greatly increases as smaller differences in calibration accuracy can be detected by standard statistical analysis. Calibration accuracy depends primarily on the standards used and the importance of validation of calibration using multiple standards from independent sources and/or standard reference materials will continue to increase.

3.5 *DTD-GC-MS* versus *solvent extraction*

Our objective was to compare the results of the three DTD-GC-MS methods to those of the solvent extraction method as the DTD-GC-MS methods would be used in place of the solvent extraction method for selected aspects of our research programme. For the spiked filter samples, this comparison could only be done for biomarkers and PAHs due to the *n*-alkane contamination.

For the spiked filter analysis, the average relative accuracies using the solvent extraction results as reference are summarised in Table 3. The *t*-test (two tailed, assumed unequal variances, $\alpha = 0.05$) was used to compare each of the DTD-GC-MS methods to the solvent extraction method. The results are summarised in Table 6.

For the diesel PM filter analysis, the average relative accuracies using the solvent extraction results as reference are summarised in Table 5. The *t*-test (two tailed, assumed unequal variances, $\alpha = 0.05$) was used to compare each of the DTD-GC-MS methods to the solvent extraction method. The results are summarised in Table 6.

Table 6. Significance test results comparing each DTD-GC-MS method with the solvent extraction method for the spiked filters and diesel PM samples.

There is no statistically significant difference (*t*-test, assumed unequal variances, $\alpha = 0.05$) between the DTD-GC-MS method and the solvent extraction method for *x* of *y* target analytes reported for the spiked filter samples.
The total number of target analytes considered by the methods is *n*.

	NRCan CETC-Ottawa	EC-ESTC	EC-AQRD
Alkanes ($n = 33$)	Not tested	Not tested	Not tested
Biomarkers ($n = 12$)	1 of 12	6 of 12	3 of 10
PAH ($n = 16$)	5 of 14	4 of 14	5 of 14

There is no statistically significant difference (*t*-test, assumed unequal variances, $\alpha = 0.05$) between the DTD-GC-MS method and the solvent extraction method for *x* of *y* target analytes reported for the diesel PM samples.
The total number of target analytes detected by the solvent extraction method is *n*.

	NRCan CETC-Ottawa	EC-ESTC	EC-AQRD
Alkanes ($n = 13$)	3 of 12	6 of 12	3 of 5
Biomarkers ($n = 11$)	4 of 8	6 of 8	0 of 6
PAH ($n = 4$)	3 of 4	2 of 4	1 of 4

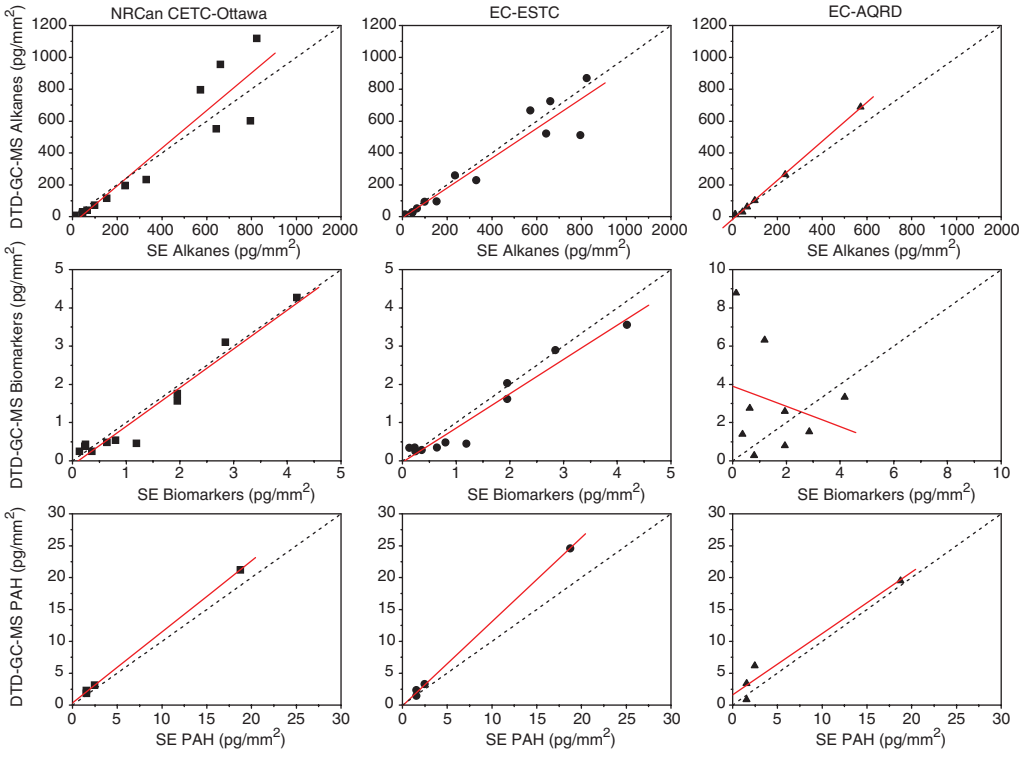


Figure 4. Correlation of thermal desorption and solvent extraction results for DPM filters. Regression lines and 1 : 1 line are shown.

For both the spiked filter samples and the diesel PM samples, less than half of the reported analytes showed no statistically significant difference between the DTD-GC-MS method and the solvent extraction method. This is a result of the comparatively good precision of the methods allowing the statistical test to detect small differences in average values. For example, the difference between the solvent extraction method and the DTD-GC-MS method of EC-ESTC for the spiked filter samples is not statistically significant for 6 of 12 biomarker compounds (Table 6); however, the average relative accuracy for the biomarkers is 9% (Table 3).

Figure 4 shows the correlation of the TD method results with the solvent extraction results by compound class. The results of linear regression analysis are summarised in Table 7. In general, the correlations were quite good, with the majority of the slopes not significantly different from unity and the majority of the intercepts not significantly different from zero. Correlation coefficients were generally greater than 0.90. There were, however, a few notable exceptions. The PAHs and the biomarker results reported by EC-AQRD differ most from the solvent extraction results, likely because of the different calibration standards and calibration curve fit used, as mentioned earlier. The results for the biomarkers are of particular concern.

The results were generally good for the PAH compounds found by both methods. Acenaphthylene was reported at a much higher concentration by the first set of solvent extraction samples, likely due to the correction for recovery (30–50%). In the second set of

Table 7. Regression analysis correlating thermal desorption results with solvent extraction results for DPM filters (±95% confidence interval).

	Slope	Intercept	R^2
Alkanes			
NRCan CETC-Ottawa	1.18 ± 0.30	-43 ± 138	0.87
EC-ESTC	0.93 ± 0.20	-6 ± 93	0.90
EC-AQRD	1.22 ± 0.06	-15 ± 16	0.99
PAHs			
NRCan CETC-Ottawa	1.11 ± 0.07	0.4 ± 0.6	0.99
EC-ESTC	1.32 ± 0.28	-0.1 ± 1	0.99
EC-AQRD	0.96 ± 0.66	1.6 ± 6	0.95
Biomarkers			
NRCan CETC-Ottawa	1.01 ± 0.17	-0.1 ± 0.3	0.95
EC-ESTC	0.90 ± 0.16	-0.04 ± 0.3	0.94
EC-AQRD	-0.53 ± 1.8	3.9 ± 3.6	0.06

Table 8. Average method detection limits (MDL, pg/mm^2) for the solvent extraction and thermal desorption methods.

	Solvent extraction	NRCan CETC-Ottawa	EC-ESTC	EC-AQRD
Alkanes	1.5	0.8	0.3	2.2
PAHs	0.8	0.8	0.1	1.8
Biomarkers	0.8	0.8	0.1	0.1
	3 times the noise on a low level standard	99% confidence interval of 6 repeat analyses of a 0.125 ng on-column standard	99% confidence interval of 7 repeat analyses of a 0.02 µg mL^{-1} standard	3 times the standard deviation of 6 repeat analyses on a low level standard

solvent extraction samples (Figure H in the Supplementary Material), acenaphthylene was reported at a concentration similar to that reported by the DTD-GC-MS methods (0.10 pg/mm^2). The results from the first and second set of solvent extraction results agreed with each other for fluoranthene, but at 1.6 pg/mm^2, this result is about 40% lower than the results from the DTD-GC-MS methods. All the other compounds detected by all methods showed no statistically significant difference.

For the biomarker compounds, the results are mixed. The higher precision of these results gave a much tighter tolerance for establishing statistical significance. There was also no consistent trend of over or underestimating the concentration by the DTD-GC-MS methods as compared to the solvent extraction method.

Lavrich and Hays [13] suggested and made use of a practical guideline for assessing statistical significance for quantification of trace organic compounds in samples such as diesel PM. They observed that while a compound may fail the statistical *t*-test, in reality, when considering the magnitudes of other sources of uncertainty that are present, that if a difference of greater than 20% is observed and the difference fails the *t*-test, then the difference can really be considered significant. If the difference is less than 20% and the *t*-test is failed, then the difference may not be significant for practical purposes.

They indicated that the factor of 20% is one that is often arbitrarily assigned to represent these other sources of measurement uncertainty. Applying this guideline to the results of this study would suggest a change in significant difference for many of the results.

The method detection limits (MDL) for the solvent extraction method and the three DTD-GC-MS methods are summarised in Table 8. Figure I in the Supplementary Material shows the MDLs by compound for each method. The DTD-GC-MS methods performed as well or better than the solvent extraction method.

4. Conclusions

Satisfactory performance (accuracy and precision) for two of the three DTD-GC-MS methods has been demonstrated for analysis of a mixture of target analytes in solution. The NRCan CETC-Ottawa and EC-ESTC methods achieved similar average relative accuracies and average relative precisions for all three classes of target analytes. The EC-AQRD method had good average relative precision but the average relative accuracies were not as good, especially for the petroleum biomarkers. This difference in accuracy was attributed to a difference in calibration strategy and a different source of calibration materials.

Higher variability was seen with all methods for the more volatile compounds (C_{10} to C_{13} *n*-alkanes and naphthalene). The *n*-alkanes were not reported by the solvent extraction method for the analysis of the unknown liquid due to contamination, but results from a subsequent set of samples analysed by an independent solvent extraction method show similar performance to the DTD-GC-MS methods for the *n*-alkanes. For the PAHs and biomarkers, the solvent extraction results also compared well to the direct analysis of the unknown liquid.

The *n*-alkane contamination was traced to the ASE cell filters. The solvents and freshly fired quartz filters were shown to be acceptably clean. A change in the solvents used for ASE cell preparation prior to sample extraction has since corrected the problem.

The results for the diesel PM filters from the three DTD-GC-MS methods gave average precisions between 7% and 21% with precisions for individual compounds generally less than 5%. Larger variations were observed when the concentration of the analyte was relatively low ($< 25\,pg/mm^2$) and for the more volatile compounds. The comparison to the solvent extraction results is complicated by the *n*-alkane contamination and differences in the DTD-GC-MS method calibrations. The average difference from the solvent extraction results ranged from 20% to 40%. Two of the methods had very similar performances while the third method was quite different. The third method was calibrated with NIST SRMs for PAHs and biomarkers and a standard prepared in-house for the *n*-alkanes. The *n*-alkane results agreed well with the solvent extraction method (19%), but the PAH and biomarker results were substantially different (82% and 907%, respectively). The ratio of highest to lowest calibration levels was 35 for the biomarkers and 7 for the PAH, each with a linear fit. The low concentration of these samples compared to the very wide calibration range suggests a linear fit may not be appropriate in this case.

This study has illustrated several challenges faced in trace organic compound analysis of particulate matter samples collected on filters. While solvent extraction methods are prone to contamination from solvents, glassware extraction cells, sample handling and/or storage, TD tubes sitting in the tray before injections and the exposure of the GC inlet

directly to ambient air during sample injection could also introduce contamination. Care must also be taken in selecting the mathematical function (e.g. linear, quadratic) used to establish the calibration curve. Depending on the range of the calibration, a linear fit may not always be the correct choice and statistical methods are available to evaluate the goodness of fit of the equation to the data. In addition, the source of the calibration standards has a large effect on the experimental results. It is important to have the analytical method validated using independent reference materials. Accuracy is improved by using a second source of standards to verify the initial calibration, but this may not always be possible, as is the case for petroleum biomarkers. The number of suppliers is very limited and there are often noticeable differences from lot-to-lot for the same supplier.

Overall, the DTD-GC-MS methods generally compare well to one another and with the conventional solvent extraction method for the analysis of *n*-alkanes, biomarkers and PAHs in PM collected on filters. The MDLs for the DTD-GC-MS methods are as good as or better than the solvent extraction method but the requirements for sample quantity are considerably less (a few micrograms for the DTD-GC-MS methods *vs.* hundreds of micrograms for the solvent extraction methods). The ability to quantify the suite of compounds examined in this paper with relatively small quantities of PM is very advantageous for airborne PM research, potentially expanding the types of samples that can be collected and the temporal resolution of the measurements.

The potential of the DTD-GC-MS technique is apparent for application to analysis of filter samples with low loading, such as those collected over shorter time intervals or in rural or pristine areas and from emission sources with highly efficient exhaust after-treatment (e.g. catalysed diesel particle filters). In addition, DTD-GC-MS does not require solvent extraction and sample clean-up so the possibility of contamination during sample preparation can be greatly reduced. Despite the fact that the DTD-GC-MS technique can be used for the analysis of PM filters with high concentration of compounds, the challenges faced when using this technique suggest that it might be better suited for PM samples with lower concentrations. The lower concentrations in ambient samples would not necessitate such a wide calibration range, there would be less material desorbed from the filter, thus lower interferences in the chromatograms and less contamination of transfer lines and inlet liners. The diesel PM samples used in this study have a large area of UCM that makes quantification of certain analytes, particularly the *n*-alkanes, difficult. Although the size of sample punch used could be reduced, this strategy has physical limits. The reduced petroleum contribution found in ambient PM samples could result in better quantification of these analytes. Therefore, future method comparisons using ambient samples are recommended.

Acknowledgements

The authors are grateful to technical assistance from Xinghua Fan, Trevor Connolly and Alison Walkey. We also acknowledge Natural Resources Canada – the Program for Energy Research and Development and Environment Canada – for financial support.

References

[1] C.A. Pope III and D.W. Dockery, J Air Waste Manage. **56**, 709 (2006).
[2] W.F. Rogge, M.A. Mazurek, L.M. Hildemann, G.R. Cass, and B.R.T. Simoneit, Atmos. Environ. A. **27**, 1309 (1993).

[3] J.J. Schauer, W.F. Rogge, L.M. Hildemann, M.A. Mazurek, G.R. Cass, and B.R.T. Simoneit, Atmos. Environ. **30**, 3837 (1996).
[4] B. Zielinska, J.C. Sagebiel, G. Harshfield, A.W. Gertler, and W.R. Pierson, Atmos. Environ. **30**, 2269 (1996).
[5] W.F. Rogge, L.M. Hildemann, M.A. Mazurek, and G.R. Cass, Environ. Sci. Technol. **32**, 13 (1998).
[6] M. Bates, P. Bruno, M. Caputi, M. Caselli, G. de Gennaro, and M. Tutino, Atmos. Environ. **42**, 6144 (2008).
[7] A.H. Falkovich and Y. Rudich, Environ. Sci. Technol. **35**, 2326 (2001).
[8] D. Waterman, B. Horsfield, F. Leistner, K. Hall, and S. Smith, Anal. Chem. **72**, 3563 (2000).
[9] S.S.H. Ho and J.Z. Yu, J. Chrom. A. **1059**, 121 (2004).
[10] T.W. Kirchstetter, C.E. Corrigan, and T. Novakov, Atmos. Environ. **35**, 1663 (2001).
[11] X. Fan, P.K.H. Lee, J.R. Brook, and S.A. Mabury, Aerosol Sci. Tech. **38**, 63 (2004).
[12] D.A. Olson and G.A. Norris, Atmos. Environ. **39**, 5437 (2005).
[13] R.J. Lavrich and M.D. Hays, Anal. Chem. **79**, 3635 (2007).
[14] R.J. Sheesley, J.J. Schauer, M. Meiritz, J.T. DeMinter, M.-S. Bae, and J.R. Turner, Aerosol Sci. and Tech. **41**, 981 (2007).
[15] G.P. Huffman, F.E. Huggins, N. Shah, R.J. Pugmire, H.L.C. Meuzelaar, M. Seehra, and A. Mannivannan, *Source Signatures of Fine Particulate Matter* (*PM2.5*) *from Petroleum Refining and Fuel Use, Progress Report*: *July 1, 1999 to December 31, 2000*, US Department of Energy Contract No. DE-AC26-99BC15220, 2000.
[16] J. Schnelle-Kreis, M. Sklorz, A. Peters, J. Cyrys, and R. Zimmermann, Atmos. Environ. **39**, 7702 (2005).
[17] M.D. Hays, N.D. Smith, J. Kinsey, Y. Dong, and P. Kariher, J. Aerosol Sci. **34**, 1061 (2003).
[18] J.F. Hamilton, P.J. Webb, A.C. Lewis, J.R. Hopkins, S. Smith, and P. Davy, Atmos. Chem. Phys. **4**, 1279, (2004).
[19] N. Ochiai, T. Ieda, K. Sasamoto, A. Fushimi, S. Hasegawa, K. Tanabe, and S. Kobayashi, J. Chrom. A. **1150**, 13 (2007).
[20] US Environmental Protection Agency Code of Federal Regulations, Schedule 40, Part 86, Control of Emissions from New and In-use Highway Vehicles and Engines, http://ecfr.gpoaccess.gov/cgi/t/text/text-idx?sid=6bd87f62e205a42d4015f0bbb9129ca3&c=ecfr&tpl=/ecfrbrowse/Title40/40cfrv18_02.tpl, 2007, cited 13 August 2008.