

Comprehensive database of Manufactured Gas Plant tars – Part A Database

Authors: Christopher Gallacher*¹, Russell Thomas², Richard Lord¹, Robert M. Kalin¹ and Chris Taylor³

¹ *Department of Civil and Env. Eng., University of Strathclyde, 75 Montrose St. Glasgow, G1 1XJ UK*

² *WSP / Parsons Brinckerhoff, Kings Orchard, 1 Queen St, Bristol, BS2 0HQ UK*

³ *National Grid Property, Warwick Technology Park, Warwick, CV34 6DA, UK*

*Corresponding Author: Christopher Gallacher, christopher.gallacher@strath.ac.uk

Abstract

RATIONALE

Coal tars are a mixture of both organic and inorganic compounds and were produced as a by-product from the manufactured gas and coke making industries. Different manufacturing processes have resulted in the production of distinctly different tar compositions. This study presents a comprehensive database of compounds produced using two-dimensional gas chromatography combined with time-of-flight mass spectrometry (GCxGC/TOFMS) analysing 16 tar samples produced by 5 distinct production processes.

METHODS

Samples of coal tar were extracted using accelerated solvent extraction (ASE) and derivatized post extraction using *N,O*-Bis(trimethylsilyl)trifluoroacetamide (BSTFA) with 1% trimethylchlorosilane (TMCS). The derivatized samples were analysed using two-dimensional gas chromatography combined with time-of-flight mass spectrometry (GCxGC/TOFMS).

RESULTS

A total of 16 tar samples originating from 5 different production processes, Low Temperature Horizontal Retorts, Horizontal Retorts, Vertical Retorts, Carbureted Water Gas and Coke Ovens, were analysed. 2369 unique compounds were detected with 948 Aromatic compounds, 196 aliphatic compounds, 380 sulphur containing compounds, 209 Oxygen containing compounds, 262 Nitrogen containing compounds and 15 mixed heterocycles. Derivatization allowed for the detection of 359 unique compounds, the majority in the form of hydroxylated polycyclic aromatic hydrocarbons, many of which would not have been detected without derivatization. Of the 2369 unique compounds detected 173 compounds were found to be present within all samples.

CONCLUSIONS

A unique comprehensive database of compounds detected within 16 tar samples from 5 different production processes was produced. The 173 compounds identified within every sample may be of particular importance from a regulatory standpoint. This initial study indicates that different production processes produce tars with different chemical signatures and can be further expanded upon by in depth analysis of the different compound types. The number of compounds presented within this database clearly demonstrates the analytical power of GCxGC/TOFMS.

Keywords: GCxGC/TOFMS, Coal Tar, Environmental Forensics, Derivatization

Introduction

Manufactured gas is a process by which solid or liquid fuels, such as coal or oil, are converted into gas, including hydrogen, carbon monoxide and methane. These processes were used between the 1806 and 1981 for the manufacture of gas in Britain, firstly for lighting and then later on for heating purposes¹. Tar is a by-product of the manufactured gas process and is produced during the carbonisation of coal, or gasification of oil. Coal tars are primarily dense non-aqueous phase liquids (DNAPLs) made up of thousands of organic and inorganic compounds. Similar tars have also been produced as a by-product of coke making. Some NAPLS produced by the Carbureted Water Gas process may also be Neutral or Light NAPLS with a specific gravity close to that of water. The organic fraction is dominated by polycyclic aromatic hydrocarbons² (PAHs) and contains many different classes of organic compounds, many present in trace quantities. Many of the organic compounds found within coal tar can be toxic, mutagenic and/or carcinogenic and have high persistence within the environment. This means that coal tar contamination forms a long-term persistent risk when it is released into the environment.

The gas industry evolved over a period of 160 year, during this time various different processes to manufacture gas were employed. Horizontal retorts were the first type of retort that was widely used, initially on all sizes of gasworks. The retorts, which was where the coal was heated in the absence of oxygen, were originally circular and made from cast iron, however due to their low durability and poor ability to withstand temperatures above 600°C, they were gradually superseded by fireclay and later silica retorts¹. Early low temperature horizontal retorts (LTHR) were heated directly by radiant heat from a shallow fuel bed of coke, beneath the furnace. This design only heated the retorts to around 600°C³ and as a result the amount of gas produced was fairly low and the decomposition of the organic compounds within the tar was limited, leading to tars more similar in nature to the parent coal⁴. The development of gas producers and regenerative furnaces combined with automated loadings systems allowed later horizontal retorts (HR) to operate more efficiently and at higher temperatures capable of exceeding 1000°C⁵ and so greater decomposition of the organic compounds within the tar occurred. The large surface area present within horizontal retorts allowed for a greater contact of the produced gases with the heated retort sides and so produced a greater degree of thermal cracking⁶.

Vertical retorts were developed in the first decade of the 20th century, these used a retort vessel which was rotated by 90° and could allow for the gas making process to be run continuously. Vertical retorts generally operated a high temperature process with temperatures similar to that of horizontal retorts however because of the design of the retort they produced significantly different tars. A temperature gradient existed within a vertical retort, especially in a continuous system where coal was continuously fed down by gravity. The tar fog generated could also escape vertically up through the coal with minimal exposure to the hot surfaces of the retort, reducing secondary decomposition. This produced a tar uniquely separate from both the LTHR and HR tars.

Two additional sources of tar are those produced by coke ovens (CO) and those produced by the carbureted water gas process (CWG). Coke oven tars are the only remaining operational producers of coal tar within the UK¹. Coke ovens were mainly located at collieries, iron and steel works, where coke was required for the smelting

process and can still be found in operation at present-day steel works. Water gas was made by injecting steam into a red hot fuel bed of coke (occasionally coal) in order to produce gas primarily consisting of hydrogen, carbon monoxide and carbon dioxide¹. The reaction was extremely endothermic, therefore water gas had to be produced by a two-step cyclical process referred to as the run and blow, the former the gas making stage and the latter the reheating cycle³. To improve the calorific and illuminating value of the water gas, the process was modified by introducing a spray of oil into the gas in the carburetor hence the name Carbureted Water Gas (CWG)⁷. A majority of the oil was fixed into the gas, any tars produced by the CWG process were consistent with oil gas tars, as the non-carbureted water gas process did not generate significant amounts of tar, when coke was used as a fuel⁸. The normal operational temperature of CWG was between 650°C and 700°C⁸ making it a relatively low temperature process.

It is estimated that over 3000 FGMP sites were present within the UK⁹. The US Environmental Protection Agency have reported that due to poor handling and storage practices in the past, contamination is likely to have occurred at up to 90% of former sites¹⁰. In 2006 it was estimated that at least 2279 former gasworks existed which had provided a public gas supply within the British Isles¹¹. This excludes privately owned gasworks and a more recent estimation of recorded public and private gasworks from 2010 estimated 3510 gasworks located within the British Isles¹².

The organic compounds found at former manufactured gas plant (FMGP) sites can be divided into three classes (pyrogenic, petrogenic and diagenetic) that indicate their mechanism of formation and potential sources¹³. Pyrogenic substances are produced from oxygen-depleted high temperature processes, whereas petrogenic substances originate from petroleum-based materials used on the site¹³. Diagenetic substances are produced by recent biological activity and are unlikely to contribute measurably to PAHs near a FMGP site¹³.

When a coal tar DNAPL is spilled into the sub-surface it will migrate vertically through the soil and underlying geology until it reaches a highly impermeable stratum and forms a long-term source of contamination. This will often lead to the DNAPL collecting within groundwater leading to the dissolution of soluble fractions of the coal tar into the groundwater. Although the rate at which contaminant mass transfer in the flowing groundwater takes place is slow the contamination may still pose a risk to human health or the environment¹⁴. D’Affonseca *et al.* (2008)¹⁴ modeled the long term degradation of coal tar contamination and found that even after 1000 years source depletion of phenanthrene was low with 89% of the original mass still remaining. The moderately and sparingly soluble composite constituents of coal tar were also predicted to have 60% and 98% of the original mass respectively remaining. These values highlight a key contamination issue of FMGP as any residual coal tar will take an extremely long time to degrade. This can form a long-term source of environmental liability and a persistent risk to human health and the environment.

Coal tar has been used as a crude chemical feedstock for the worldwide fine-chemical industry, such as production of dyes or creosote for wood treatment⁹. This adds environmental forensic complications, as it is possible to encounter coal-tar contamination in a broad variety of sites other than FMGPs. By the late 1800’s, coal tar was used in the synthesis of a wide array of industrial materials and consumer products, including: dyes, perfumes, explosives and pharmaceuticals³. Tars that were

not sold to refiners may have been landfilled or disposed of in open pits⁷ resulting in a potential for tar contamination at many former landfill sites. Coal tar and other FMGP by-products and wastes were also deposited on FMGP site if space was available and ground raising was required¹.

The use of GCxGC allows for the separation of unresolved complex mixtures (UCM), often referred to as a big “hump” in the GC chromatogram¹⁵, which cannot be separated using traditional GM/MS analysis. Because UCMs are believed to consist of many thousands of compounds, traditional GM/MS is simply not capable of providing sufficient resolution leaving most UCM hydrocarbons unidentified. The combination of the GCxGC with the TOFMS (time-of-flight mass spectrometer) allows for the identification of compounds within the UCM that previously would be unidentifiable within the “hump” and most likely missed using GM/MS. The use of GCxGC therefore allows for the separation of coal tar without the need for a lengthy separation process.

While a large amount of research surrounding coal tar is been published the production processes used to produce the tar are often not reported, as they are likely unknown. A comprehensive database of the compounds found in tars from various different production processes is therefore required. The following study presents a database of compounds found in 16 different tars produced by 5 distinct production processes including: Horizontal Retort, Vertical Retorts, Low Temperature Horizontal Retorts, Carbureted Water Gas plants and Coke Ovens.

Materials and Methods

Samples:

A total of 16 tars samples, coming from 5 different tar production processes, were extracted and analysed. The different processes and sample numbers are listed below (full site information for each sample can be found in the supplementary information):

Low Temperature Horizontal Retort (LTHR): DNAPL009 (D9L) and DNAPL016 (D16L)

Vertical Retort (VR): DNAPL002-006 (D2-D6V) and DNAPL020 (D20V)

Horizontal Retort (HR): DNAPL007 (D7H), DNAPL008 (D8H), DNAPL010 (D10H) and DNAPL017 (D17H)

Carbureted Water Gas (CWG): DNAPL013 (D13C) and DNAPL014 (D14C)

Coke Oven (CO): DNAPL018 (D18CO) and DNAPL019 (D19CO)

Methods:

All solvents used were of analytical grade purchased from Fisher Scientific (Loughborough, U.K.) and D₁₀-phenanthrene, which was used as an injection standard, was purchased from Sigma-Aldrich (Gillingham, U.K.). D₈-naphthalene, D₁₀-fluorene, D₁₀-fluoranthene and D₁₀-pyrene, which were used as recovery standards, was purchased from Sigma-Aldrich (Gillingham, U.K.). *N,O*-Bis(trimethylsilyl)trifluoroacetamide (BSTFA) with 1% trimethylchlorosilane

(TMCS) was purchased from Sigma-Aldrich (Gillingham, U.K.). Silica was purchased from Fisher Scientific (Loughborough, U.K.)

Extraction was performed using an Accelerated Solvent Extraction system (ASE 350 Dionex, Camberley, UK) using 10 mL stainless steel extraction cells. Approximately 0.5g of tar was mixed with an equal amount of diatomaceous earth in a 1:1 ratio. Prior to extraction the samples were spiked a recovery standard. Extraction cells were lined with 2 Dionex glass fibre filter papers and packed with 3g of silica gel 60 deactivated with 10% water. The sample mixture was then loaded into the cells and any residue was recovered with excess diatomaceous earth. Dichloromethane was used as the extracting solvent for all extractions. ASE was performed at 100°C and 10 MPa, using one dynamic (7 min) and two static (5 min each) extractions. A flush volume of 150% and purge time of 60 s was used. The extracts were concentrated to 1 mL using a Büchi Syncore Analyst (Oldham, U.K). The extracts were then made up to exactly 10 mL using *n*-hexane. A 1 mL aliquot was then transferred to an auto sampler vial prior to analysis and spiked with D₁₀-phenanthrene. All samples were derivatized using 100ul of BSTFA with 1% TMCS placed in an oven at 70°C for 1 hour.

GCxGC TOFMS analysis was performed using a Leco Pegasus 4D (St. Joseph, Michigan, USA) time of flight mass spectrometer, connected to an Agilent 7890A (Wilmington, Delaware, USA) gas chromatograph equipped with a LECO thermal modulator. The TOF ion source temperature was 200 °C and the mass range 45 and 500u was scanned at a rate of 200 spectra/second. The detector voltage was set at 1700 V with an electron ionisation energy of 70 eV.

All standards and extracts were analysed with the following primary oven temperature programme: 60°C isotherm for 2 minute, then ramp at 10°C/min to 110°C, then ramp at 3°C/min to 310 °C, and isothermal at 310°C for 15 minutes. The secondary oven and modulator temperatures were programmed at a 10 °C offset relative to the primary oven. The modulation period was 6 seconds with a 1.3 second hot pulse time and a cool time of 1.7 seconds. The injection port temperature was set to 250 °C and set to split injection with a split ratio of 50 and an injection volume of 1µL. Helium was used as the carrier gas, with a flow rate of 1.0 mL/min.

The reversed polarity column set that was used comprised of a mid-polarity TR-50 MS supplied by Thermo Scientific (Hertfordshire, U.K.) (30 m × 0.25 mm i.d. × 0.25 µm film thickness) as the primary column and a non-polar Rtx-5SilMS supplied by Thames Restek (Buckinghamshire, U.K.) (1.5 m × 0.25 mm i.d. m × 0.25 µm film thickness) as the secondary column, connected via a Thames Restek Press-tight connector.

The chromatograms from each sample were processed using Leco ChromaTOF software (Version 4.50.8.0) to search for, identify and align all peaks with a signal-to-noise ratio greater than 10. As it would be impractical to purchase standards of every compound present within the samples any concentration comparisons are relative comparisons of the peak area for each compound adjusted for sample weight and recovery. Only direct like for like compound comparison was performed, as the detector response will remain equivalent.

Results and Discussion

Database:

A total of 5 different production processes are represented within the coal tar database, these included: Low Temperature Horizontal Retorts (LTHR), Horizontal Retorts (HR), Vertical Retorts (VR), Carbureted Water Gas (CWG) and Coke Ovens (CO). These 5 production processes cover a wide range of types of coal tars manufactured, although it should be noted that the list is not exhaustive and other production processes have been used. Figure 1 shows the overall composition of each sample as well as the total number of each compound class present within the samples. The aliphatic class includes *n*-alkanes, *n*-alkenes, branched alkanes/alkenes and alkyl cyclohexane/pentanes. The PAH class includes all parent and alkyl PAHs. The derivatizable class includes all hydroxylated PAHs (such as phenolics) as well as hydroxylated PASHs, which has been previously reported in coal liquids¹⁶. The mixed heterocycles include heterocyclic compounds with more than one element substituted within the ring such as oxygen-sulphur. Thienobenzofurans are an example of a group of mixed heterocycles detected for which no literature reporting their presence within coal tar could be found. Nitrogen-sulphur mixed heterocycles such as azabenzothiophene were also detected, which have been previously reported in coal liquids¹⁷ and anthracene oil¹⁸. The PANH class includes all nitrogen containing compounds present in the database, with the exception of mixed heterocycles. The PAOH class contains all oxygen containing compounds, with the exception of hydroxylated compounds and mixed heterocycles. The PASH class contains all sulphur analogues of polycyclic aromatic hydrocarbons, with the exception of hydroxylated sulphur compounds and mixed heterocycles.

A total of 16 coal tar samples were analysed with 2369 unique compounds identified with the largest group being PAHs containing 948 compounds. The database also includes 380 PASHs, 209 PAOHs, 262 PANHs and 15 mixed heterocycles. Finally a total of 196 aliphatic compounds were identified as well as 359 derivatizable compounds. A full list of the compounds found in all samples including retention times can be found within the supplementary information.

Of the 2369 compounds detected 173 compounds, shown in Table 1, were found to be present in all samples. The majority of these compounds were PAHs with 126 PAHs being found in all samples. Alkyl PAHs dominate the list with a total of 77 alkyl PAHs detected. While many alkyl PAHs were detected within every sample many other isomer were also detected within the database as a whole. For example while 4 C₂-fluorene were detected within every sample 23 C₂-fluorenes were present in the overall database. Due to potential limitations with the use of the EPA16 PAHs, which is commonly used for the assessment of contaminated sites, the EPA34 list¹⁹ was created in order to include several groups of alkyl PAHs that may be of interest. Of the EPA34 PAHs 30 were detected within every sample with only C₃-fluorene, C₃-benzanthracene/chrysene, C₄-benzanthracene/chrysene and C₄-phenanthrene/anthracene isomers not being detected in every sample. As the EPA34 is not a list of 34 individual compounds and in fact is made up of 18 parent PAHs combined with 16 groups of prominent C₁ to C₄ alkyl PAHs a total of 70 individual compounds that fall within the EPA34 were detected within every sample. A total of 202 compounds detected that fall within the EPA34 that were also detected in multiple samples. While C₃-fluorene isomers were not detected within every sample they were detected within every sample except D10H. C₃-Benzanthracene/chrysene isomers were detected within 12 of the 16 tar samples, whereas C₄-

benzanthracene/chrysene was detected within only 5. Finally C₄-phenanthrene/anthracene isomers were detected within only 8 of the 16 tar samples.

Several other parent PAHs and groups of alkyl PAHs were detected within every sample that do not fall within the range of the EPA34 and could be of potential forensics and legislative interest when dealing with coal tar contaminated sites. The lowest molecular weight PAH that was detected within every sample was styrene (*m/z* 104) with the highest molecular weight PAH detected within every sample being dibenzochrysene (*m/z* 276). Several potential groups of compounds of interest were identified such as C₂-fluoranthene/pyrene, 5 isomer of which were detected within every sample, as C₁-fluoranthene/pyrene isomers are included within the EPA34 but the C₂ isomers are not. Benzo[a]pyrene is included within the EPA34 and is often used for the assessing the carcinogenic and mutagenic risk of a contaminated site²⁰. However alkylated benzo[a]pyrene isomers are not included either within the EPA34 or risk assessments and for example both 6-methyl benzo[a]pyrene and 10-methyl benzo[a]pyrene are more mutagenic than benzo[a]pyrene itself²¹. A total of 7 potential isomers of benzo[a]pyrene were identified within every sample, although the specific parent PAH isomer could not be determined. This suggests that the identification of specific alkylated benzo[a]pyrene isomers within coal tars may be of great potential scientific, and legislative, interest. The parent PAHs indene, indane, cyclopenta[def]phenanthrene, cyclopenta[ghi]perylene and indeno[2,1,7-cde]pyrene were detected within every sample but are not included within the EPA34 and therefore might also be of potential interest. Dibenzopyrene isomers, C₂-benzo[x]fluoranthene and C₁-benzo[ghi]perylene isomers were also detected within every sample, although no single isomer of each was present within every sample, and so are not included within the 173 compounds detected within every sample.

A total of 6 PANHs were detected within every sample with 5 heterocyclic PANHs as well as 1-naphthalenecarbonitrile, which is the only compound detected within every sample containing active nitrogen. Carbazole as well as 2 of its C₁ and 1 of its C₂ alkylated isomers were detected within every sample with benzo[def]carbazole making up the remaining PANH. Several other PANHs were detected in the majority of the samples with for example C₃-pyridine detected within every sample except for samples D2V and D3V. Other PANHs detected in the majority of the samples include benzo[c]carbazole and benzo[b]carbazole, that were detected in every sample except for D3V, and benzo[h]quinoline that was detected in every sample except D2V. As both samples D2V and D3V are in the form of DNAPL contaminated soils these compounds might be expected to be contained within all pure coal tar DNAPLs. It should be noted that while 2 C₁-carbazole isomers eluting at 40.1, 2.040 and 41.5, 2.015 respectively were detected within every sample the isomers detected within D10H were below the signal to noise ratio of 10 and therefore would not be quantifiable using standards. This suggests that while these compounds may be likely present within all coal tar DNAPLs they will not always be present in quantifiable concentrations.

A total of 10 PAOHs were detected within every sample with 9 heterocyclic PAOHs and a single aromatic ketone in the form of benz[de]anthracen-7-one. Benzofuran as well as 1 C₁ and 3 C₂ benzofuran isomers were detected within every sample as well as benzo[b]naphtha[2,3-d]furan, benzo[b]naphtha[2,1-d]furan and benzo[b]naphtha[1,2-d]furan, with benzo[k,l]xanthene detected in every sample

except D2V. A single C₁ isomers of benzonaphthofuran was detected within every sample as well as C₂ isomers being detected in every sample except D2V. This suggests that both benzo[k,l]xanthene and C₂ benzonaphthofuran isomers would be expected to be detected in purer coal tar DNAPL samples. While no single phenolic compound was found to be present within every sample phenol was detected within every sample except for D20V. 4-Hydroxyfluorene and C₃-phenol isomers were also detected within every tar sample except for D2V and D3V. This suggests that 4-hydroxyfluorene and C₃-phenol isomers would be expected to be present in purer forms of coal tar DNAPL, although this is not the case for phenol as D20V is in the form of pure DNAPL taken from a tar tank.

All of the 31 PASHs detected within every sample are in the form of heterocyclic PASHs, with alkylated isomers dominant. Dibenzothiophene, naphtho[1,2-b]thiophene, naphtho[2,1-b]thiophene and 2 phenanthro/phenaleno thiophene isomers were detected within every sample with the remaining 26 PASHs present in the form of alkylated isomers. Naphtho[1,2-b]thiophene is the only proven mutagenic 3 ring PASH²² and was detected within every sample. The most mutagenic 4 ring PASHs, phenanthro[3,4-b]thiophene²² was found to be present within 10 of the 16 samples and was found to be present within all HR tars suggesting it may be a compound of concern when dealing with HR tars. Benzo[2,3]phenanthro[4,5-bcd]thiophene is an extremely potent mutagen, more potent than benzo[a]pyrene²², and was detected within 9 of the 16 sample including both CWG and CO tars suggesting that it may be a compound of concern when dealing with these two tar types.

The most abundant PASH isomers are alkylated benzo[b]thiophenes with 3 C₁, 4 C₂ and 5 C₃ benzo[b]thiophene isomers detected within every sample. Benzo[b]thiophene itself was not detected within every sample however but was detected within every sample except for D2V suggesting that it would be expected to be detected within the majority of coal tar DNAPLs. Alkylated dibenzothiophenes are the next most abundant group of PASHs with 5 C₁ and 4 C₂ isomers detected within every sample. C₃-Dibenzothiophene isomers were detected within every sample except for D16L, thus making them another group of compounds of potential interest. Alkylated thiophenes were not detected within every sample however C₂-thiophene isomers were detected within every sample except D6V and C₃-thiophene isomers were detected in every sample except the 2 CO tars. This suggests that C₃-thiophenes may be potential compounds of interest for manufactured gas tars rather than those tars produced by coke ovens.

All of the PASHs detected within every sample were heterocyclic PASHs with a single sulfur present within the aromatic ring. However a benzodithiophene isomer, which contains 2 sulfurs within the aromatic ring, was detected within every sample except for D2V. Due to the nature of sample D2V, DNAPL contaminated soil, this suggests that benzodithiophene isomers would be expected to be detected within purer tar DNAPLs. The highest molecular weight PASHs detected within every sample were 2 C₁-benzonaphthothiophene isomers, however higher molecular weight PASHs were detected within the majority of the tar samples with both pyreno[1,2-b]thiophene and pyreno[2,1-b]thiophene detected within every sample except D3V. C₁-Benzonaphthothiophene isomers were also detected within every sample except D16L. Finally the highest molecular weight PASHs detected in the majority of the samples were benzophenanthrothiophene isomers, which were detected in every

sample except D16L. This suggests that all of these compounds would be expected to be detected within the majority of tar DNAPLs, and particularly within pure tar DNAPL.

Only 2 aliphatic compounds were detected within all samples, the C₁₁ and C₁₂ n-alkanes, while C₁₃ to C₁₆ n-alkanes occurred in every sample except D18CO (coke oven). Both coke oven samples were produced from the same feedstock coal, and sampled on the same day, suggesting that the higher temperatures used to produce D18CO, relative to D19CO, results in the destruction of n-alkanes. n-Alkenes were also detected in every sample except the coke oven tars ranging between C₁₁ and C₁₄, with only a single C₁₀ alkene detected in D19CO and no alkenes detected in D18CO. This suggests that alkenes are destroyed, or at the very least not created, during the production of coke. Branched alkanes ranging from C₁₁ to C₁₃ were detected within every sample except the coke oven tars, with only a single branched alkane present within sample D19CO. This suggests that while the presence or absence of alkanes, up to C₁₆, may not be used to distinguish manufactured gas tars from coke oven tars the presence or absence of branched alkanes and alkenes may be able to be used.

Comparison of General trends between Tar types:

The 16 tar samples present within this study fall into 3 rough sample composition types with D7H, D8H, D10H, D17H, D9L, D14C, C18CO, D19CO and D20V being present in the form of pure DNAPL. Samples D2V, D3V, D4V, D5V, D6V and D16L are all present in the form of DNAPL saturated soil. Finally sample D13C is present in the form of contaminated water from a groundwater contamination plume. This means that the samples of pure DNAPL are likely more indicative of their original tars, with different degrees of potential weathering. The tar samples from contaminated soils will have a significantly higher dilution factor than pure DNAPL and an increased possibility for the loss of original components within the tar. The tar sample taken from a groundwater plume, D13C, will be both more dilute than the original DNAPL as well as having likely lost at least some of its more water soluble organics.

Figure 1 shows the total number of individual compounds detected within each sample type broken down into chemical class. When comparing the pure DNAPL samples, and therefore excluding the contaminated soils and water sample, several interesting trends are present. The LTHR tar D9L displays the most aliphatic (136) and a large number of derivatized (203) compounds relative to the other tar samples suggesting that a high number of aliphatic and phenolic compounds is indicative of a LTHR tar. The VR tar D20V, which is the only pure VR DNAPL, displays both a large number of aliphatic (124) and derivatized (231) compounds as well as a lower number of individual PANHs (135) compared to other pure DNAPLs, with the exception of the CWG tar C14C. The CWG tar D14C, which is the only pure CWG DNAPL, contains the highest number of PAHs (729) within the database as well as the highest number of PASHs with alkylated isomers dominant, suggesting that a large number of these compounds should be detected within CWG tars. D14C also contains the lowest number of PANHs (102) of the pure DNAPLs and a relatively small amount of PAOHs, with the exception of D10H. The two CO tars both contain similar numbers of compounds with similar overall trends. The CO tars both contain a low number of aliphatic compounds with 2 in D18CO and 9 present in D19CO. The CO tars contain significantly less individual PAHs and PASHs, with the exception of

D10H, relative to the other pure DNAPLs and less PAOHs relative to the pure VR tars, LTHR tars and HR tars, with the exception of D10HR and the CWG tars

The overall total number of compounds within the HR tars varies widely with D9H containing 1517, D17H containing 1306, D7H containing 1290 and D10H containing only 872. All of the HR tars are in the form of pure DNAPL and so the large variance in the number of compounds cannot be explained by environmental dilution. The variations may be down to variations in the original feedstocks used to produce the tars. D10H is significantly different from the other HR tars with an unusually small number of PAHs (334) relative to all the tar samples and PASHs (143) relative to the other HR tars samples. D10H has the 2nd lowest number of PAHs within the database sample with only D16L, which is a contaminated soil, containing less individual PAHs. When the remaining HR tars are taken as whole several general trends are present. The remaining HR tars contain less aliphatic compounds than the other tar types, with the exception of CO tars, and less derivatized compounds than the LTHR and VR tars but more than the CO and CWG tars. The HR tars contain similar numbers of PAHs to LTHR and VR tars but less than CWG and more than CO tars.

The total number of compounds within the remaining non-pure DNAPL tar samples varies widely even within tar samples taken from the same site. D2V to D6V were all taken from the same site, although D2V was taken from a different part of the site from D3V to D6V. D2V contains the lowest number of overall compounds of any sample within the database with 650 individual compounds detected. The remaining VR tars have similar total compounds although with a high degree of variability of the compound types and D4 containing the most overall compounds with 1118. The different number of compounds detected may be down to where the samples were taken. D3V, D4V and D5V were all taken from an in filled tar tank so it is possible that D4V contains more of the original tar relative to the other samples from the same site. D16L contains only 672 individual compounds and this is likely because the sample is a tar-contaminated soil from a site that ceased operation in 1946. This means that the sample has spent a significant amount of time within the environment and has likely undergone significant weathering. Finally sample D13C contains 1150 individual compounds despite the fact the sample comes from contaminated groundwater.

Conclusion

A unique comprehensive database of compounds detected within 16 tars produced by 5 different production processes was produced. In total 2373 unique compounds were detected with 173 compounds detected within all samples. The 173 compounds detected within every sample are of potential forensics importance for source apportionment and also to aid current risk assessments relating to coal tar contaminated sites. These currently focus on a limited number of compounds, usually the 16 priority PAHs in the UK. Several of the 173 compounds detected within every sample may also be of potential interest from a toxicological standpoint. The results clearly demonstrated that different production processes produce distinctly different coal tars with a degree of variability within the tars produced by the same processes. This initial study clearly demonstrates the power of GCxGC/TOFMS for the forensic analysis of environmental samples and can be further expanded on by more detailed examination of the individual compound types.

References

1. R. A. P. Thomas. *The History and Operation of Gasworks (Manufactured Gas Plants) in Britain*. CL:AIRE, **2014**.
2. P. S. Birak, C. T. Miller. Dense non-aqueous phase liquids at former manufactured gas plants: challenges to modeling and remediation. *J. Contam. Hydrol.* **2009**, *105*, 81–98.
3. S. M. Harkins, R. S. Truesdale, R. Hill, P. Hoffman & S. Winters. *U.S. Production of Manufactured Gases: Assessment of Past Disposal Practices*. Prepared by Research Triangle Institute. Prepared for Hazardous Waste Engineering Research Laboratory, U.S. Environmental Protection Agency, Cincinnati, OH; EPA 600-2-88-012. **1988**.
4. R. P. Soule. *Examination of Low-Temperature Coal Tar*. PhD thesis. Columbia University, New York. **1922**.
5. W. J. Butterfield. *The Chemistry of Gas Manufacture: A Practical handbook of the production, purification, and testing of illuminating and fuel gas, and on the bye-products of gas manufacture*. Edition 3, Volume 1 Materials and Processes London: Charles Griffin and Company, Limited; Exeter Street, Strand, **1904**.
6. S. Young. *Distillation Principles and Processes*. MacMillan and Company, London, **1922**.
7. B. L. Murphy, T. Sparaci, W. J. Shields. Manufactured gas plants — processes, historical development, and key issues in insurance coverage disputes. *Enviro. Forensics.* **2005**, *6*, 161–173.
8. M. J. Hamper. Manufactured Gas History and Processes. *Enviro. Forensics/* **2006**, *7*, 55–64.
9. L. A. McGregor, C. Gauchotte-Lindsay, N. Nic Daéid, R. Thoma, R. M. Kalin. Multivariate statistical methods for the environmental forensic classification of coal tars from former manufactured gas plants. *Environ. Sci. Technol.* **2012**, *46*, 3744–52.
10. U.S. EPA. *Cleaning up the Nation's Waste Sites: Markets and Technology Trends*. United States Environmental Protection Agency, Office of Solid Waste and Emergency Response, Washington, DC. EPA 542-R-04-015, **2004**
11. J. Horne. How Many Gasworks Were There?. *Historic Gas Times*, **2006**, *46*, 2a
12. R. Thomas. *Gasworks Sites in the UK and Ireland (sites recorded to Date)*, **2010**.
13. D. Saber, D. Mauro, T. Sirivedhin. Environmental Forensics Investigation in Sediments near a Former Manufactured Gas Plant Site. *Enviro. Forensics.* **2006**, *7*, 65–75.
14. F. M. D'Affonseca, P. Blum, M. Finkel, R. Melzer, P. Grathwohl. Field scale characterization and modeling of contaminant release from a coal tar source zone. *J. Contam. Hydrol.* **2008**, *102*, 120–39.
15. C. Tran, Tin C, G. A. Logan, E. Grosjean, D. Ryan, P. J. Marriott. Use of comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry for the characterization of biodegradation and unresolved

- complex mixtures in petroleum. *Geochim. Cosmochim. Acta.* **2010**, 74, 6468–6484.
16. M. Nishioka, M. L. Lee, H. Kudo, D. R. Muchiri, L. J. Baldwin, S. Pakray, J. G. Stuart, R. N. Castle. Determination of Hydroxylated Thiophenic Compounds in Coal Liquid. *Anal. Chem.* **1985**, 1330, 1984–1987.
 17. M. Nishioka, R. M. Campbell, W. R. West, P. Smith, G. M. Booth, M. L. Lee, H. Kudo, R. N. Castle. Determination of Aminodibenzothiophenes in a Coal Liquid. *Anal. Chem.* **1985**, 57, 1868-1871.
 18. P. Burchill, A. Herod, P. Ernest. Investigation of nitrogen compounds in coal tar products. 2. Basic fractions. *Fuel.* **1983**, 62, 20–29.
 19. H. P. H. Arp, N. A. Azzolina, G. Cornelissen, S. B. Hawthorne. Predicting pore water EPA-34 PAH concentrations and toxicity in pyrogenic-impacted sediments using pyrene content. *Environ. Sci. Technol.* **2011**, 45, 5139–5146.
 20. A. L. Juhasz, R. Naidu. Bioremediation of high molecular weight polycyclic aromatic hydrocarbons : a review of the microbial degradation of benzo [a] pyrene. *Int. Biodeterior. Biodegrad.* **2000**, 45, 57-88.
 21. L. M. Calle, P. D. Sullivan, M. D. Nettleman, I. J. Ocasio, J. Blazyk, J. Jollick. Antioxidants and the mutagenicity of benzo(a)pyrene and some derivatives. *Biochem. Biophys. Res. Commun.* **1978**, 85, 351–356.
 22. J. Jacob. *Sulfur analogues of polycyclic aromatic hydrocarbons (thiaarenes)* Cambridge monographs on Cancer Research. Cambridge University Press, Cambridge. **1990**.

Acknowledgements

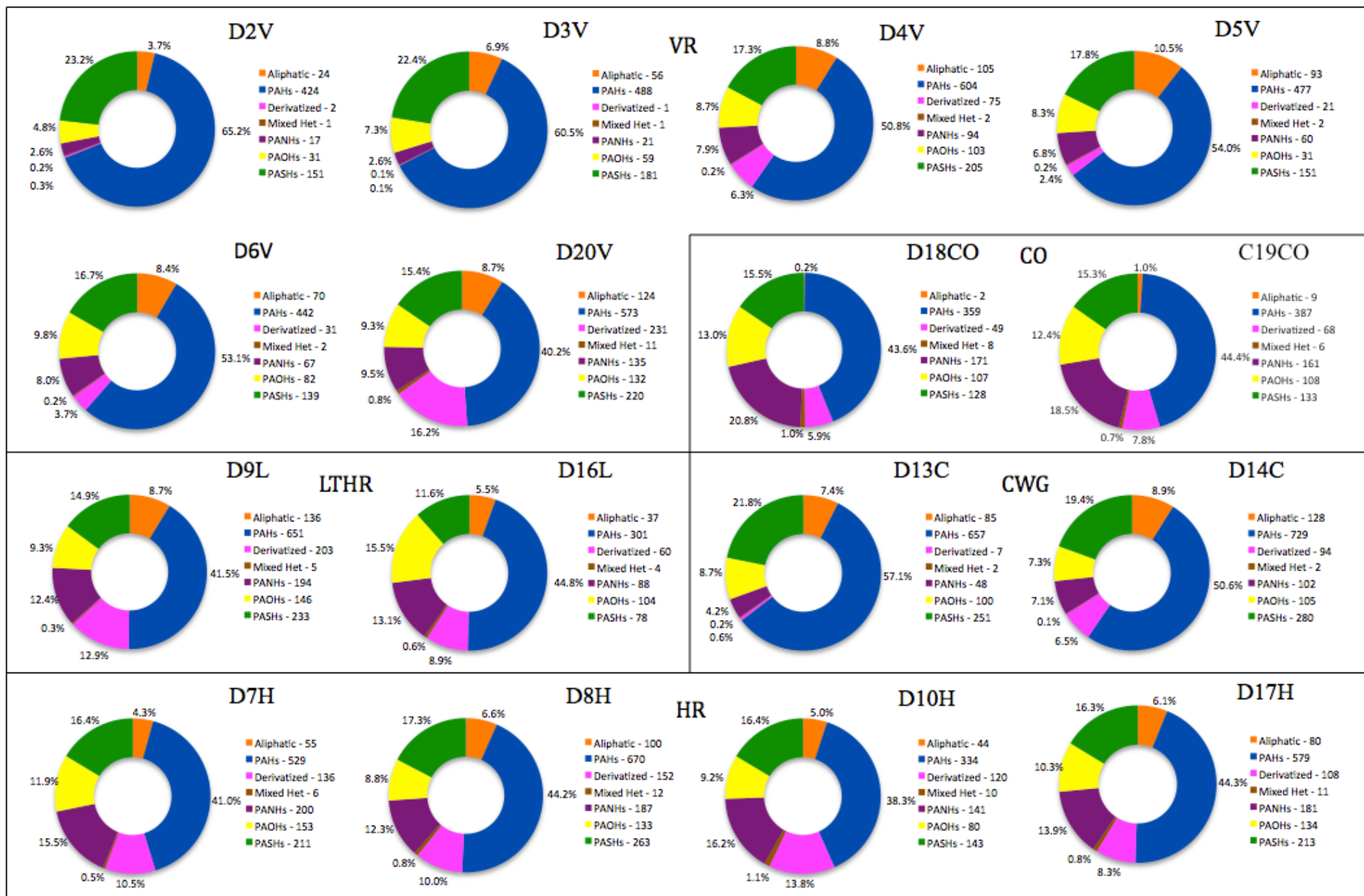
We thank the Scottish Funding Council (SFC) Glasgow Research Partnership in Engineering, the University of Strathclyde, WSP Parsons Brinckerhoff and National Grid property for funding support. The authors declare no competing financial interests.

Attached:

Figure 1 - Total number of compounds per sample (HR = Horizontal Retort, VR = Vertical Retort, LTHR = Low Temperature Horizontal Retort, CWG = Carbureted Water Gas, CO = Coke Oven)

Compound	m/z	No	Formula	Compound	m/z	No	Formula	Compound	m/z	No	Formula
C11 n-alkane	57	1	C ₁₁ H ₂₂	C1-Phenanthrene/Anthracene *EPA34	192	6	C ₁₅ H ₁₂	Dibenz(a,h)anthracene *EPA16 + EPA34	278	1	C ₂₂ H ₁₄
C12 n-alkane	57	1	C ₁₂ H ₂₄	C2-Flourene *EPA34	194	4	C ₁₅ H ₁₄	m/z 278 PAH C22H14	278	2	C ₂₂ H ₁₄
Styrene	104	1	C ₈ H ₈	Fluoranthene *EPA16 + EPA34	202	1	C ₁₆ H ₁₀	Benzo[g,h,i]perylene *EPA16 + EPA34	276	1	C ₂₂ H ₁₂
Indene	116	1	C ₉ H ₈	Pyrene *EPA16 + EPA34	202	1	C ₁₆ H ₁₀	Dibenzochrysene	276	1	C ₂₂ H ₁₂
Indane	118	1	C ₉ H ₁₀	Acenaphenanthrylene/Aceanthrylene	202	1	C ₁₆ H ₁₀	1-Naphthalenecarbonitrile	153	1	C ₁₁ H ₇ N
C3-Benzene	120	3	C ₉ H ₁₂	C2-Phenanthrene/Anthracene *EPA34	206	6	C ₁₆ H ₁₄	Carbazole	167	1	C ₁₂ H ₉ N
Naphthalene *EPA16 + EPA34	128	1	C ₁₀ H ₈	C4-Biphenyl	210	1	C ₁₆ H ₁₈	C1-Carbazole	181	2	C ₁₄ H ₉ N
Dihydronaphthalene	130	2	C ₁₀ H ₁₀	C1-Fluoranthene/Pyrene *EPA34	216	5	C ₁₇ H ₁₂	Benzo[def]carbazole	191	1	C ₁₄ H ₉ N
Methyl Indene	130	1	C ₁₀ H ₁₀	C3-Phenanthrene/Anthracene *EPA34	220	3	C ₁₇ H ₁₆	C2-Carbazole	195	1	C ₁₄ H ₁₃ N
C4-Benzene 1DB	132	1	C ₁₀ H ₁₂	Benzo[ghi]fluoranthene	226	1	C ₁₈ H ₁₀	Benzofuran	118	1	C ₈ H ₆ O
Methyl Indan	132	1	C ₁₀ H ₁₂	Benzo[c]Phenanthrene	228	1	C ₁₈ H ₁₂	C1-Benzofuran	132	1	C ₉ H ₈ O
C4-Benzene	134	2	C ₁₀ H ₁₄	Benzo[a]anthracene *EPA16 + EPA34	228	1	C ₁₈ H ₁₂	C2-Benzofuran	146	3	C ₁₀ H ₁₀ O
C1-Naphthalene *EPA34	142	2	C ₁₁ H ₁₀	Chrysene *EPA16 + EPA34	228	1	C ₁₈ H ₁₂	Benzo[b]naphtho[2,3-d]furan	218	1	C ₁₆ H ₁₀ O
Acenaphthylene *EPA16 + EPA34	152	1	C ₁₂ H ₈	Triphenylene/Naphacene	228	1	C ₁₈ H ₁₂	Benzo[b]naphtho[2,1-d]furan	218	1	C ₁₆ H ₁₀ O
Acenaphthene *EPA16 + EPA34	154	1	C ₁₂ H ₁₀	C2-Fluoranthene/Pyrene	230	5	C ₁₈ H ₁₄	Benzo[b]naphtho[1,2-d]furan	218	1	C ₁₆ H ₁₀ O
Biphenyl	154	1	C ₁₂ H ₁₀	9H-Cyclopenta[a]pyrene isomers	240	3	C ₁₉ H ₁₂	7H-Benz[de]antracen-7-one	230	1	C ₁₇ H ₁₀ O
C2-Naphthalene *EPA34	156	8	C ₁₂ H ₁₂	C1-Benzanthracenes/Chrysene *EPA34	242	4	C ₁₉ H ₁₄	C1-Benzonaphthofuran	232	1	C ₁₇ H ₁₂ O
Fluorene *EPA16 + EPA34	166	1	C ₁₃ H ₁₀	C1-Dihydro. Benz[a]anthracene	244	1	C ₁₉ H ₁₆	C1-Benzothiophene	148	3	C ₉ H ₈ S
1H-Phenylene	166	1	C ₁₃ H ₁₀	Benzo[b]fluoranthene *EPA16 + EPA34	252	1	C ₂₀ H ₁₂	Phenyl Thiophene	160	2	C ₁₀ H ₈ S
Benz[x]indene isomers	166	3	C ₁₃ H ₁₀	Benzo[k]fluoranthene *EPA16 + EPA34	252	1	C ₂₀ H ₁₂	C2-Benzothiophene	162	4	C ₁₀ H ₁₀ S
Diphenylmethane	167	1	C ₁₃ H ₁₂	Benzo[e]pyrene *REACH + EPA34	252	1	C ₂₀ H ₁₂	C3-Benzothiophene	176	5	C ₁₁ H ₁₂ S
C1-Biphenyl	168	2	C ₁₃ H ₁₂	Benzo[a]pyrene *EPA16 + EPA34	252	1	C ₂₀ H ₁₂	Dibenzothiophene	184	1	C ₁₂ H ₈ S
C3-Naphthalene 1DB	168	1	C ₁₃ H ₁₂	Perylene *EPA34	252	1	C ₂₀ H ₁₂	Naphtho[1,2-b]thiophene	184	1	C ₁₂ H ₈ S
C3-Naphthalene *EPA34	170	5	C ₁₃ H ₁₄	Binaphthalene	254	2	C ₂₀ H ₁₄	Naphtho[2,1-b]thiophene	184	1	C ₁₂ H ₈ S
Phenanthrene *EPA16 + EPA34	178	1	C ₁₄ H ₁₀	C2-Benzanthracenes/Chrysene *EPA34	256	1	C ₂₀ H ₁₆	C1-Dibenzothiophene	198	5	C ₁₃ H ₁₀ S
Anthracene *EPA16 + EPA34	178	1	C ₁₄ H ₁₀	Cyclopenta[ghi]perylene	264	1	C ₂₁ H ₁₂	Phenanthro/Phenylene Thiophene	208	2	C ₁₄ H ₈ S
C1-Fluorene *EPA34	180	6	C ₁₄ H ₁₂	11H-Indeno[2,1,7-cde]pyrene	264	1	C ₂₁ H ₁₂	Phenyl Benzothiophene	210	1	C ₁₄ H ₁₀ S
C4-Naphthalene *EPA34	184	5	C ₁₄ H ₁₆	C1-Benzo[x]fluoranthene isomers	266	7	C ₂₁ H ₁₄	C2-Dibenzothiophene	212	4	C ₁₄ H ₁₂ S
Cyclopenta[def]phenanthrene	190	1	C ₁₅ H ₁₀	Indeno[1,2,3-cd]pyrene *EPA16 + EPA34	276	1	C ₂₂ H ₁₂	C1-Benzonaphthothiophene isomers	248	2	C ₁₇ H ₁₂ S

Table 1 – Compounds present in all database samples *A full list including retention times has been included in the supplementary data



Supplementary information: Site History

DNAPL001-006 (D2-D6) samples were taken from a site that had been used to manufacture gas between 1836 and 1971. The site had used a wide range of process from low temperature horizontal retorts, a high temperature horizontal retort and later a vertical retort. D2 was taken from a borehole near a gasholder on the southern boundary of the site. D3, D4 and D5 were taken directly from a former tar tank on the opposite side of the site. This tar tank had been built at the same time as the vertical retort plant had been constructed. D6 was taken from a borehole located near to the tar tank containing samples D3, D4 and D5. All the samples were produced by the Vertical Retort (VR) process. D2, D3 and D6 are Sandy soil samples heavily contaminated with tar. D4 and 005 are both waterlogged soils that are heavily contaminated with tar.

The site initially operated simple horizontal retorts until the gasworks was redeveloped in 1878 and expanded in 1912 to include five continuous vertical retort beds. This was further extended in the 1920s with the addition of an additional five continuous vertical retorts. From the early 1930s until 1952 the plant remained at capacity, with various improvements to increase efficiency, such as the addition of two water gas plants. The site was expanded again in 1952 with the construction of a purification plant, additional vertical retorts and the development of a storage area for primary flash distillate (a type of light petroleum distillate). A CWG (carbureted water gas) plant and microsimplex oil gasification plant operated on the site and so there is the potential for there to be traces of CWG and oil gas tar within the samples.

DNAPL007 (D7) was taken from a site that operated between 1856 and 1969 and initially produced low temperature horizontal retort (LTHR) gas. However, this site later switched to higher temperature horizontal retort gas and the tar sample came from this process rather than the earlier LTHR process. The sample is in the form of pure DNAPL (dense non-aqueous phase liquids) taken at the base of an underground gasholder tank during a remediation project in 2009.

DNAPL008 (D8) was collected from a FMGP (former manufactured gas plant), which operated between 1856 and 1971, and it may contain traces of CWG tar, as the site produced CWG in the later stages of its operation. The majority of the tar would have been produced by the HR process and this was shown in the statistical analysis of the tars⁹. The sample is in the form of pure DNAPL in water. The tar sample was from a former 250,000 gallon tank associated with the horizontal retort house and was taken from a skimmer pump, which was being used to empty the tar tanks during a remediation process

DNAPL009 (D9) was obtained from a site that ceased production in 1953, with an unknown initial operation date. The site was manufacturing gas using a vertical retorts when it closed, but could also have used a horizontal retort, probably a LTHR, or carbureted water gas. The CWG plant was commissioned in 1902, with vertical retorts being constructed in 1913 and 1931. Statistical analysis showed the samples to have been derived from the LTHR process⁹. The sample is in the form of pure DNAPL in water and was taken from inside a former tar well during a remediation process. The sample was taken within a clean glass laboratory sample of 250mL.

DNAPL010 (D10) was obtained from a FMGP, which operated between 1849 and 1981 and mainly operated horizontal retorts, originally using LTHR and later switching to a fired HR process *circa* 1890. The tar samples were collected from a circular concrete tar tank known to be associated with the later HR process. The sample is in the form of pure DNAPL in water. The site also operated a CWG plant so it is possible that CWG tar may also be present within the sample.

DNAPL013 (D13) operated from 1885 until the 1970s. The site manufactured gas using a horizontal retort until 1939, taking a bulk gas supply from another source. It did, however, continue to maintain and use its CWG plant to supply peak demand. The sample of liquid NAPL was removed from a groundwater NAPL plume beneath the site by a skimmer pump.

DNAPL014 (D14) was from the same site as DNAPL017 and was taken from an infilled former gasholder and later a tar emulsion storage tank, used for the storage of NAPL emulsions produced by the CWG process, and it was pure DNAPL. This sample was obtained from a remediation system, which used total fluids pumping

DNAPL016 (D16) was obtained from a FMGP, which operated between 1854 and 1946, and was a typical small country town FMGP site. The site is known to have operated hand-charged directly fired horizontal retorts, which would have produced a tar with LTHR characteristics, as shown in the statistical analysis ⁹. The sample was recovered from the core barrel of a pioneer rotary drilling rig and is in the form of a tar-saturated soil.

DNAPL017 (D17) was taken from the same site as DNAPL014 (D14), a site that operated from 1868 to 1976. The site operated a HR, coke ovens, CWG, a gas reforming plant and a by-products works. Samples D14 and D17 were obtained from an area of the site where both the CWG and the HR had operated. D17 was taken from a tar tank that would have been associated with the HR process.

DNAPL018 (D18) and DNAPL019 (D19) were taken from coke ovens that were still in operation at a steelworks. Both samples were taken on the day of coke production with D18 coming from an oven that has been in operation since the 1970s. D19 was taken from a coke oven that has been in operation since the 1930s. Both samples are in the form of pure DNAPL. The coal used in the manufacture of these tars was a blend of up to 12 non-British imported crushed coals.

DNAPL020 (D20) was taken from a site that operated from 1885 to 1975. The site initially used horizontal retorts until the 1920s before then operating a VR plant and a CWG plant. The coal tar sample was derived from the VR process. The sample is in the form of pure DNAPL taken from a site drain. As the site also operated a CWG plant, CWG tar may be present in trace amounts.