ELECTRIC POWER
RESEARCH INSTITUTE

Examination of the Sources of Polycyclic Aromatic Hydrocarbon (PAH) in Urban Background Soil

Examination of the Sources of Polycyclic Aromatic Hydrocarbon (PAH) in Urban Background Soil

1015558

Interim Report, December 2008

EPRI Project Manager
J. Lingle

ELECTRIC POWER RESEARCH INSTITUTE
3420 Hillview Avenue, Palo Alto, California 94304-1338 ▪ PO Box 10412, Palo Alto, California 94303-0813 ▪ USA
800.313.3774 ▪ 650.855.2121 ▪ askepri@epri.com ▪ www.epri.com

DISCLAIMER OF WARRANTIES AND LIMITATION OF LIABILITIES

THIS DOCUMENT WAS PREPARED BY THE ORGANIZATION(S) NAMED BELOW AS AN ACCOUNT OF WORK SPONSORED OR COSPONSORED BY THE ELECTRIC POWER RESEARCH INSTITUTE, INC. (EPRI). NEITHER EPRI, ANY MEMBER OF EPRI, ANY COSPONSOR, THE ORGANIZATION(S) BELOW, NOR ANY PERSON ACTING ON BEHALF OF ANY OF THEM:

(A) MAKES ANY WARRANTY OR REPRESENTATION WHATSOEVER, EXPRESS OR IMPLIED, (I) WITH RESPECT TO THE USE OF ANY INFORMATION, APPARATUS, METHOD, PROCESS, OR SIMILAR ITEM DISCLOSED IN THIS DOCUMENT, INCLUDING MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, OR (II) THAT SUCH USE DOES NOT INFRINGE ON OR INTERFERE WITH PRIVATELY OWNED RIGHTS, INCLUDING ANY PARTY'S INTELLECTUAL PROPERTY, OR (III) THAT THIS DOCUMENT IS SUITABLE TO ANY PARTICULAR USER'S CIRCUMSTANCE; OR

(B) ASSUMES RESPONSIBILITY FOR ANY DAMAGES OR OTHER LIABILITY WHATSOEVER (INCLUDING ANY CONSEQUENTIAL DAMAGES, EVEN IF EPRI OR ANY EPRI REPRESENTATIVE HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES) RESULTING FROM YOUR SELECTION OR USE OF THIS DOCUMENT OR ANY INFORMATION, APPARATUS, METHOD, PROCESS, OR SIMILAR ITEM DISCLOSED IN THIS DOCUMENT.

ORGANIZATION(S) THAT PREPARED THIS DOCUMENT

Meta Environmental

NOTE

For further information about EPRI, call the EPRI Customer Assistance Center at 800.313.3774 or e-mail askepri@epri.com.

Electric Power Research Institute, EPRI, and TOGETHER…SHAPING THE FUTURE OF ELECTRICITY are registered service marks of the Electric Power Research Institute, Inc.

Copyright © 2008 Electric Power Research Institute, Inc. All rights reserved.

CITATIONS

This report was prepared by

Meta Environmental
49 Clarendon Street
Watertown, MA 02172

Principal Investigators
D. Mauro
J. Roush

This report describes research sponsored by the Electric Power Research Institute (EPRI).

The report is a corporate document that should be cited in the literature in the following manner:

Examination of the Sources of Polycyclic Aromatic Hydrocarbon (PAH) in Urban Background Soil. EPRI, Palo Alto, CA: 2008. 1015558.

PRODUCT DESCRIPTION

Between 2000 and 2005, EPRI collected several hundred soil samples from urban background locations in three States and analyzed them for 17 polycyclic aromatic hydrocarbon compounds (PAHs). This report presents the initial results of efforts to extract additional chemical data from the urban background PAH database and to explore those data using chemical forensic methods for statistical properties and trends.

Results and Findings

All additional PAHs and related compounds extracted from the data files were lognormally distributed as were the original 17 PAHs. The distributions of low molecular weight compounds were more highly skewed because of the greater numbers of non-detects. In contrast, the distributions of PAH ratios were more normally distributed. For most ratios, a few outliers, high or low, were the main deviations from normality. Therefore, simple summary statistics of the PAH ratios could be used to determine if there were statistically significant differences between urban background and samples from former manufactured gas plant (MGP) sites.

Some PAH ratios or combinations of PAH ratios were found to be consistently different when comparing background PAHs and MGP-derived PAHs. Some PAH ratios showed little discriminating power. Of particular value, the ratio of fluoranthene/pyrene (Fl/Py) falls between about 0.5 and 0.8 for tar-like material (TLM) from former carbureted water gas (CWG) MGPs and falls between about 0.9 and 1.3 for urban background. However, the Fl/Py ratio for former coal carbonization (CC) MGP tars also falls between about 0.9 and 1.3. Therefore, other ratios are needed. For example, the ratio of benz(a)anthracene/chrysene (BAA/CHR) is greater in CC MGP TLM and coal tar and creosote from coking operations than it is urban background. Other PAH ratios with some discriminating power include retene/benzo(a)pyrene and the sum of fluoranthene and pyrene verses total HPAHs. The use of these ratios for PAH source attribution has been reported by many authors.

Based on a comparison of the general composition and PAH ratios of a large number of urban background samples to those of several samples from former MGP sites, this study indicates that much, if not all, of the general urban background PAH compounds in soil analyzed in previous EPRI studies are from non-MGP sources.

Challenges and Objectives

It is well known that there are many sources of PAHs in urban areas, of which former manufactured gas plants (MGPs) was one. MGP site managers often must determine whether PAHs found in samples off the site originated at the site. This can be a complex problem given the many potential PAH sources within urban areas. The objective of this study was to explore a large database of chemical data collected from shallow surface soils in urban areas by EPRI for

potential diagnostic and indicator compounds, statistical distributions, and trends that discriminate between former MGP impacts and general urban background.

Applications, Values, and Use

This study provides MGP site owners with a high quality, robust data set to use as a benchmark or baseline when trying to determine whether PAHs found in urban soil near their sites originated with the former MGP operation. If they suspect that the PAHs originated with another operation or are part of the general urban background, then MGP site owners have several environmental forensic methods available with which to attribute the PAHs to those sources as discussed in several other EPRI reports and the open literature.

EPRI Perspective

PAH compounds are ubiquitous, therefore understanding the nature and source of the contaminants is important and challenging. Previous EPRI investigations of background PAH contamination levels in urban soils focused on the statistical distributions of individual PAH compounds. This "data mining" effort takes a significant step further by using a variety of chemical forensic techniques to reexamine the potential sources of PAH contamination in those urban soils tested. In the process, a much better understanding of the forensic techniques available and limitations has resulted. This report should be of use to MGP site managers and others investigating background urban PAH soil contamination.

Approach

Surface soil samples from 318 sites in 29 population centers in three states were collected and analyzed for PAHs. The samples were analyzed for wide-range hydrocarbon fingerprint by gas chromatography with flame ionization detection (GC/FID) and 43 PAHs and alkylated PAHs by GC with mass spectrometric detection (GC/MS). Using the electronic GC/MS data files from the previous analyses of the soil and QC samples, selected alkylated PAHs were identified and their respective concentrations calculated using the original calibration data. In addition, several compounds of potential diagnostic value were included in the data re-processing. Then, a subset of the data files were examined for the presence of certain petroleum biomarker compounds. The data were compiled into a spreadsheet and basic descriptive statistics calculated. Also, selected compound ratios were calculated and examined graphically.

Keywords

MGP
PAH
Environmental forensics

CONTENTS

1 INTRODUCTION AND BACKGROUND 1-1
Introduction 1-1
Definition of Background 1-1
Sources of PAHs in the Environment 1-2
Composition of Pyrogenic and Petrogenic Materials 1-4
Description of Chemical Fingerprinting Methodology 1-4
GC/FID Fingerprinting 1-5
Extended PAH Profiles (EPPs) by GC/MS 1-5
Data Mining 1-7
Objectives 1-8

2 METHODS AND MATERIALS 2-1
Source and Collection of Samples 2-1
Sample Collection, Preparation and Analysis Methods 2-1
Definitions 2-1
Study Area 2-2
Site Selection 2-2
Sample Locations 2-3
Sample Collection and Compositing 2-3
Sample Preparation and Analysis 2-3
Illinois and New York Samples 2-4
Pennsylvania Samples 2-4
Quality Assurance and Quality Control 2-4
Extraction of Alkylated PAHs and other Target Compounds from the Data Set 2-5
Potential Diagnostic Compounds Not Expected to Be in the Data Set 2-5
Cleanups 2-6
Detection limits 2-6
Data Analysis 2-6

***3* RESULTS AND DISCUSSION.....3-1**
Site and Area Land Use Types.....3-1
Summary of Physical Sample Characteristics 3-2
Sample Results for PAHs 3-2
Summary Statistics for PAH Results from All Sites 3-2
Summary Statistics for PAH Results by Site and Area Uses.....3-4
PAH Sources and Data Trends 3-6
Total Extractable Hydrocarbons.....3-6
Petroleum Biomarkers 3-12
PAH Ratios 3-13
Fluoranthene/pyrene 3-14
Benz(a)anthracene/Chrysene.....3-17
Benzofluorenes/Methylpyrenes 3-19
Other PAH Ratios.....3-20
Multivariate Exploratory Analysis 3-26

***4* SUMMARY 4-1**

***5* REFERENCES AND BIBLIOGRAPHY.....5-1**

***A* TARGET ANALYTE LIST A-1**

LIST OF FIGURES

Figure 3-1 Normal Probability Plot for Phenanthrene Concentrations in Surface Soil (N=318) 3-4
Figure 3-2 Normal Probability Plot for Total PAH Concentrations in Surface Soil (N=318) 3-5
Figure 3-3 GC/FID Chromatograms Showing Typical Urban Background Signatures 3-8
Figure 3-4 Severely Weathered TLM-Impacted Soil Sample from a Former MGP Site Showing Substantial Loss of LPAHs 3-9
Figure 3-5 Normal Probability Plot for the Ratio LPAH/HPAH 3-11
Figure 3-6 Frequency Histogram for the Ratio of Total Priority Pollutant PAHs to Total PAHs(43) 3-12
Figure 3-7 Distribution of Fluoranthene/Pyrene Ratios in Urban Background Soil 3-16
Figure 3-8 Comparison of Fluoranthene/Pyrene Ratios in Urban Background to Tar-Like Materials from Former CWG MGP Sites 3-17
Figure 3-9 Comparison of Urban Background Soil and Severely Weathered MGP Residuals in Soil at a Former MGP Site 3-18
Figure 3-10 Benz(a)anthracene/Chysene Ratios of Urban Background Soil Samples and TLM 3-18
Figure 3-11 Double Ratio Plot of Benz(a)anthracene/Chrysene verse Fluoranthene/Pyrene 3-19
Figure 3-12 Double Ratio Plot: Retene/Benzo(a)pyrene verses Benz(a)anthracene/Chrysene for Urban Background Soil and TLM Samples 3-22
Figure 3-13 Benzo(e)pyrene/Benzo(a)pyrene Ratio in Urban Background Soil and TLM Samples 3-23
Figure 3-14 Trialkylated-dibenzothiophenes/Trialkylated-phenanthrenes verses Dialkylated-dibenzothiophenes/Dialkylated-phenanthrenes 3-24
Figure 3-15 Ratio of Benzo(b)fluoranthene/Benzo(k)fluoranthene verses Total PAHs in Urban Background Soil and Selected TLM Samples 3-25
Figure 3-16 The Ratio of Fluoranthene+Pyrene/HPAHs verses Total PP PAHs in Urban Background Soils and TLM Samples 3-26
Figure 3-17 Principal Components Analysis of Normalized Background PAH Data Set 3-27

LIST OF TABLES

Table 3-1 Summary Statistics for All Samples in the EPRI Background PAH Study 3-3
Table 3-2 Summary Statistics by Site Uses for Benzo(a)pyrene 3-5
Table 3-3 Summary Statistics by Area Uses for Benzo(a)pyrene 3-6
Table 3-4 Descriptive Statistics for TOC and Total PAH Characteristics in Urban Background Surface Soil 3-10
Table 3-5 Comparison of Median and Upper and Lower Quartiles of Characteristic Ratios for Urban Background Soil Samples to Tars and Petroleum Products 3-14
Table 3-6 Descriptive Statistics for Selected PAH Ratios in Urban Background Surface Soil 3-15

1
INTRODUCTION AND BACKGROUND

Introduction

Polycyclic aromatic hydrocarbons (PAHs) are ubiquitous in the environment. EPRI recently completed a study of the distribution of PAHs in surface soil in urban areas in three states. In that study, EPRI contractors collected over 300 shallow soil samples from random locations in 30 urban areas and analyzed them by EPA Method 8270 for aromatic hydrocarbons. Data were provided on the concentrations and statistics for the 16 Priority Pollutant PAHs defined under the U.S. EPA Clean Water Act plus 2-methylnaphthalene (EPRI 2002, EPRI 2003a, EPRI 2004, EPRI 2008). These are the generally regulated PAH compounds and the ones typically analyzed for as part of environmental site assessments.

However, studies have shown that environmental samples can contain hundreds of PAHs and that the distributions of those PAHs can provide clues to their sources (Boehm 2006). This EPRI report summarizes the initial efforts to extract additional PAH data from the EPRI background PAHs database and to evaluate whether those data can be used to determine PAH sources.

A general summary of the nature and sources of PAHs in the environment is provided in the next subsections followed by a brief introduction to the methods used to measure PAHs and evaluate PAH data for source identification purposes. Section 2 of this report reviews the sample collection and analysis methods used in the original EPRI background PAH studies and methods for extracting the supplemental PAH data from the data files. Section 3 presents and summarizes the supplemental data.

Definition of Background

The definitions of "background" PAHs in the literature and in current regulations are neither concise nor consistent. For example, the Massachusetts Department of Environmental Protection (MADEP) defines background as "considered to be those levels of organic hazardous material (OHM) which would exist in the absence of the disposal site" (MADEP 1992). According to USEPA (1999b), a background level is:

"…the concentration of hazardous substance that provides a defensible reference point with which to evaluate whether or not a release from the site has occurred. The background level should be reflective of the concentration of the hazardous substance in the medium of concern for the environmental setting on or near a site. Background level does not necessarily represent pre-release conditions, nor conditions in the absence of influence from source(s) at the site. Background level may or may not be less than the detection limit, but if it is greater than the

detection limit, it should account for variability in local concentrations. Background levels need not be established by chemical analysis."

Thus, the term "background" as currently used in the literature and regulatory documents does not have a clear or consistent meaning. The definition of "background" should consider sample location with respect to known or suspected PAH sources. For example, background PAH levels generally are very different for urban and rural locations, with higher concentrations measured in urban settings where a greater frequency of PAH sources are present.

For the purpose of this report, the term, background PAHs, will refer to the PAHs in the environment from everyday life (atmospheric deposition from natural and manmade sources, general urban fill, road dust, and so on) and not directly affected by any single known source of PAH-containing material.

Sources of PAHs in the Environment

It is well known that urban PAHs originate from many potential sources, both petrogenic and pyrogenic. In general, petrogenic substances include crude oil and refined crude oil products such as gasoline, heating oil, asphalt, and coal. Pyrogenic substances are complex mixtures of primarily hydrocarbons produced from organic matter subjected to high temperatures but with insufficient oxygen for complete combustion. Pyrogenic PAHs are produced by fires, internal combustion engines, and furnaces. They also are formed when coke or gas are produced from coal or oil. Coal-tar based products, such as roofing, pavement sealers, waterproofing, pesticides, and some shampoos contain pyrogenic PAHs (Boehm 2006).

Aromatic hydrocarbons include monocyclic aromatic hydrocarbons (MAHs) such as benzene, toluene, xylenes, and alkylated benzenes, and PAHs such as naphthalene, phenanthrene, and pyrene. MAHs and PAHs originate from many sources and exist at many sites. This section briefly reviews the sources of PAHs in urban soils and sediment.

Crude petroleum, many of its refined products, coal, coal tar, and many coal tar products consist primarily of hydrocarbons. Hydrocarbons are organic molecules that are made up of only carbon and hydrogen atoms. Some simple hydrocarbons include hexane and benzene. There are several types of hydrocarbons that are commonly grouped by similar chemical structures, such as alkanes, cyclic alkanes, and aromatic hydrocarbons.

PAHs are one group of hydrocarbons that are present at high relative amounts in crude oil, coal, coal tar, and many of their products. In environmental forensic chemistry and geochemistry, PAHs are placed in subgroups according to their origins. These groups include diagenic, petrogenic.

Diagenic PAHs are formed by natural biological and chemical processes at ambient temperatures. When present, these PAHs are found at very low concentrations, typically less than a few parts per billion in soil.

Petrogenic PAHs are formed at relatively low temperatures, during the maturation of crude oil and coal. Both crude oil and coal contain hundreds of different PAH compounds, including many that are the subject of environmental investigations and are regulated.

Petrogenic PAHs have been released into urban environments from numerous anthropogenic sources over the past two centuries. For example, it has been a common practice to spray roads with oil to manage dust. Asphalt is produced from petroleum and the small particles that are created as roads wear away contain PAHs. Cars and trucks drip fuels and lubricating oils that contain petrogenic PAHs. Many industries have stored and ultimately spilled petroleum products that range from gasoline to heavy oils. Further, the potential impacts from coal cannot be ignored. For many years, residential and commercial buildings were heated with coal and small amounts of coal and coal dust accumulated wherever coal was handled. All of these sources of petrogenic PAHs, and many others, contributed to a pervasive background of PAHs in urban settings. Because many releases occurred years ago at unanticipated locations and because soil was moved around as the urban environment expanded and was modified for various uses, it is difficult to predict where and at what levels petrogenic PAHs might be found.

Finally, PAHs are formed whenever organic substances are exposed to high temperatures under low oxygen or no oxygen conditions in a process called pyrolysis. Pyrolytic processes occur intentionally, such as in the destructive distillation of coal into coke and coal tar, or the thermal cracking of petroleum residuals into lighter hydrocarbons and oil tar. Similar processes occur unintentionally, such as the incomplete combustion of motor fuels in cars and trucks, the incomplete combustion of wood in forest fires and fireplaces, and the incomplete combustion of coal and fuel oils in heating systems. These processes occur at temperatures that range from about 350°C to more than 1200°C, and their products are called pyrogenic.

Like petrogenic PAHs, pyrogenic PAHs have been released into urban environments from numerous sources. These include some obvious sources, such as building fires and industrial smoke stacks. They also include less obvious sources, such as debris from coal tar-treated roofing and building materials. The incomplete combustion of gasoline and diesel fuel in cars, trucks, and buses produces substantial amounts of pyrogenic PAHs that attach to small particles in the atmosphere and accumulate along roadsides. Any industry that utilized high temperatures in their operation probably produced PAHs. These included such industries as foundries, steel mills, coke plants, smelters, and others. Similar to petrogenic PAHs, pyrogenic PAHs accumulated in soil and are found throughout all urban areas.

Much modern gasoline is unusual in that it contains both petrogenic substances (the light distillate of crude oil) and pyrogenic substances (the light hydrocarbons from thermal cracking of oil). For the purposes of this report, all motor gasoline is considered petrogenic.

Studies conducted by META Environmental, Inc. and others indicate that much shallow urban soil contains a complex mixture of petrogenic and pyrogenic PAHs.

Composition of Pyrogenic and Petrogenic Materials

Both pyrogenic and petrogenic sources of PAHs have been found to contain hundreds of individual PAH compounds in generally predictable patterns. For example, it is known that the temperature of formation of PAHs largely determines the distribution of the various parent and alkylated PAHs. The structures of parent or unsubstituted PAHs are various combinations of fused benzene rings. Naphthalene is a simple 2-ring parent PAH. In contrast, alkylated PAHs have an alkane or cyclic alkane substitution on a parent PAH. For example, 2-methylnaphthalene and 2-ethylnaphthalene are simple alkylated PAHs. The possible combinations of alkyl-substitutions on parent PAHs is very large accounting for the many hundreds of compounds found (Boehm 2006).

Generally, petrogenic and pyrogenic substances are comprised of many of the same compounds. It is the distribution of those compounds, or relative concentrations, that distinguish them. For example, the high temperature, low oxygen conditions under which pyrogenic PAHs are formed favor the formation of parent PAHs over substituted PAHs. Hence, parent PAHs are abundant in pyrogenic substances, not alkylated PAHs. In contrast, alkylated PAHs (and other non-aromatic hydrocarbons) are abundant in petrogenic substances, not parent PAHs.

Further, and of particular importance to environmental forensic chemistry and geochemistry is the fact that petrogenic and pyrogenic substances from different sources can have measurably different amounts of some PAHs. For example, crude oils from different reservoirs can exhibit notably different ratios of trialkylated dibenzothiophenes to trialkylated phenanthrenes. Similarly, the ratio of dialkylated chrysene to chrysene varies among certain pyrogenic sources. Consequently, the determination of PAH profiles forms an important component of environmental forensic studies where hydrocarbon releases, either petrogenic or pyrogenic, are known or suspected to be involved.

Finally, in addition to PAHs, pyrogenic and petrogenic substances contain various amounts of paraffinic hydrocarbons, olefinic hydrocarbons, naphthenic hydrocarbons, and other types of compounds. The amounts of sulfur-, oxygen-, or nitrogen-containing compounds can vary from source to source and can be used to compare samples. Therefore, in environmental forensic studies, the presence and relative amounts of these compounds also is used to identify the nature and source of hydrocarbon-based materials in environmental samples.

As discussed above, PAHs in urban background originate from both petrogenic and pyrogenic sources. At some locations, a particular petrogenic source might contribute the majority of the PAHs. Alternatively, at other locations, a particular pyrogenic source might be the major source of the PAHs. Perhaps more commonly, several petrogenic and pyrogenic sources have contributed PAHs to soil in urban locations. It is this potential for multiple PAH sources that makes source identification in urban soil challenging.

Description of Chemical Fingerprinting Methodology

PAHs commonly form the basis for source attribution and allocation at sites involving petrogenic or pyrogenic materials. Studies have shown that the pattern of PAHs clearly distinguishes

petrogenic from pyrogenic substances and can be used to identify and classify petrogenic or pyrogenic substances of different origins. For example, ASTM Method D 5739-95 is the method used extensively by the U.S. Coast Guard to determine the source of oil spilled in public waterways. That method relies on the determination of selected PAHs in oil, soil, or water samples by gas chromatography with mass spectrometric detection (GC/MS) and the use of the qualitative patterns and quantitative ratios of those PAHs to determine which oil samples have a common origin. Similarly, work by META Environmental, Inc. (META) has shown that the same methodology can be used to identify the sources of PAHs at former MGP sites, coke plants, tar refineries and wood treating facilities (EPRI 2000a).

GC/FID Fingerprinting

All soil samples in this study were analyzed by gas chromatography with flame ionization detection (GC/FID). With GC/FID, organic compounds in a sample are vaporized and then separated in a long, narrow fused silica capillary column. Separation follows boiling point approximately with the most volatile compounds exiting the column first followed by increasingly less volatile compounds. Therefore, certain refined petroleum products, generated by the distillation of crude oil and which differ in their boiling point ranges, are distinguishable by where they appear on a chromatogram. Once they exit the column, the compounds are detected using the flame ionization technique. As the compounds exit and are detected, their responses are recorded and shown as peaks on a continuous plot. The height and area of a peak are proportional to the concentration of that compound in the sample. When done in a controlled and reproducible manner, the GC/FID method produces a "fingerprint" of a sample where the presence and relative amounts of the compounds are immediately visible as peaks of varying height appearing at different times.

Source identification using GC/FID is mostly qualitatively applied. An experienced chemist examines the chromatograms, compares them to those of reference materials, and makes a judgment regarding the nature and source of the contamination in the sample. The chemist might go "peak-by-peak" looking for similarities and differences, comparing peak ratios, and looking for indicator compounds.

For some samples, GC/FID fingerprinting is accurate and sufficient. However, the reliability of GC/FID fingerprinting decreases when multiple sources are present in a sample and when the sample composition becomes extensively altered by environmental weathering processes. As discussed above, shallow urban background soil at many locations likely contains PAHs from multiple sources. Further, because the shallow soil is exposed to weather, substantial alteration of the original PAH compositions are expected from processes such as rain water leaching, evaporation/volatilization, and biodegradation. Other testing methods, particularly GC/MS, provide complementary information for source identification under these conditions.

Extended PAH Profiles (EPPs) by GC/MS

In the initial phases of the EPRI background PAH study, all of the soil samples were analyzed by GC/MS for the 16 U.S. EPA priority pollutant PAHs plus 2-methylnaphthalene. As part of this

current data mining study, an expanded list of parent and alkylated PAHs was extracted from the original raw data files and is the subject of this report.

Separation was accomplished with gas chromatography using a method similar to the GC/FID method discussed previously. However, in GC/MS, once compounds exit the column, they are detected using a mass spectrometer. In the mass spectrometer, the molecules of each compound are ionized at high temperature and vacuum. The ionic fragments are unstable and fragment into smaller ions. The ions are then counted and the mass spectrum recorded. Thus, the mass spectrum for a compound is the pattern of ionic fragments that forms when that compound is ionized. Mass spectra vary widely and are characteristic of their source compound. For example, the mass spectrum of hexane is very different from the mass spectrum of benzene even though both compounds contain six carbon atoms plus hydrogen atoms.

In GC/MS, one obtains both a chromatogram of peaks and additional compound-specific information in the mass spectrum. When executed in a controlled and reproducible manner, the GC/MS method produces multiple "fingerprints" of a sample when specific fragment ions are isolated.

GC/MS is utilized in two general ways in environmental forensic chemistry. First, samples are analyzed under the conditions required by various standard methods, particularly EPA Methods 8260 and 8270 (U.S. EPA SW-846). The concentrations of certain target compounds are determined and the mass spectrum of each peak in the chromatogram is generated and stored. These mass spectra can be used to identify non-target compounds or to generate extracted ion current profiles (EICPs). Second, various specialty methods are utilized where the GC/MS operating conditions are setup to measure only certain groups of compounds. For example, the method described in 40 CFR Subchapter J Part 300 Subpart L Appendix C for PAHs, alkylated PAHs, and biomarkers is used extensively in oil spill and UST release analyses. This method is similar to ASTM Method D 5739-95, "Standard Practice for Oil Spill Source Identification by Gas Chromatography and Positive Ion Electron Impact Low Resolution Mass Spectrometry."

GC/MS data are used both qualitatively and quantitatively. An experienced chemist examines the chromatograms, compares them to those of reference materials, and makes judgments regarding the nature and source of the contamination in the sample. The chemist might go "peak-by-peak" looking for similarities and differences, comparing peak ratios, and looking for indicator compounds. This process is described in detail in ASTM Method D 5739-95.

GC/MS data are more commonly used quantitatively by calculating the concentrations of selected compounds, by comparing peak area ratios, or by applying chemometric or pattern recognition techniques to the raw or adjusted data. These data analysis methods are used extensively with extended PAH profiles (PAHs and alkylated PAHs) and with biomarker compound data. Various degrees of statistical confidence can be achieved by examining chemical concentrations and compound ratios or patterns from multiple samples and replicate samples. This characteristic of GC/MS quantitative data is particularly valuable when assessing the degree of similarity or difference between samples, particularly when multiple sources of hydrocarbons are present in the sample or when environmental weathering has altered the original distributions of hydrocarbons.

Finally, the mass spectra of selected compounds also can be examined to determine whether any diagnostic or indicator chemicals are present in the sample. For example, the PAH retene (1-methyl-7-isopropylphenanthrene) is present in significant concentrations in coal, but at much lower concentrations in coal tar or petroleum products. Thus, the ratio of retene to chrysene can be used to explore whether coal fines are present in a soil sample and to explain some of the hydrocarbon patterns observed at sites where coal was used extensively. Further, unknown compounds can be identified and their presence used as clues to the source(s) of the chemicals.

The GC/MS data in this study were reported and utilized both qualitatively and quantitatively. First, the concentrations of the 16 priority pollutant PAHs plus 2-methylnaphthalene were calculated. Subsequent to that, the concentrations of alkylated PAHs and selected marker compounds were determined. Appendix 1 provides a list of compounds determined in this study. These data were utilized to estimate contaminant levels in samples, to generate graphs and compare compound ratios.

The GC/MS data also are used qualitatively by generating extracted ion current profiles (EICPs) for selected compounds and compound groups of forensic value. For example, the EICPs for selected "biomarker" compounds including normal alkanes, isoprenoid hydrocarbons, alkylcyclohexanes, triterpanes and steranes were generated. These compound groups are commonly used in hydrocarbon source identifications and weathering evaluations. For example, the estimated boiling point range of a refined petroleum product, as indicated by the location of the alkanes and unresolved complex mixture (UCM) on the chromatogram, can be used to determine whether the material is kerosene, diesel, No. 6 fuel oil, or some other product. Similarly, triterpanes and steranes are known to be present in crude oils and some refined petroleum products, but not found in coke oven tars and rarely found in MGP tars. Therefore, the presence of triterpanes and steranes is monitored to confirm and refine the petrogenic verses pyrogenic assessment conducted with the PAH profiles.

This project continues the process of examining background PAH concentrations by extracting the alkylated PAH profiles and concentrations and other selected hydrocarbons from the existing EPRI database and examining the distributions of those concentrations and selected compound ratios.

Data Mining

Data Mining is an analytic process designed to explore data (usually large amounts of data) in search of consistent patterns and/or systematic relationships among variables, and then to validate the findings by applying the detected patterns to new subsets of data. The ultimate goal of data mining is prediction. The process of data mining consists of three stages: (1) the initial exploration, (2) model building or pattern identification with validation/verification, and (3) deployment (i.e., the application of the model to new data in order to generate predictions). Stage 2 (model building) and Stage 3 (deployment) activities were not conducted during this phase of the EPRI project.

The data mining steps completed for this phase of the study were as follows.

- Pre-Stage 1: Data Extraction. The raw instrument data files, including calibration and quality control samples, generated over several years, were assembled and organized to facilitate easy processing and/or query. The data files were reviewed by an experienced GC/MS analyst for various characteristics and compounds determined to be of potential forensic value. These characteristics and compounds included selected alkylated PAHs, heterocyclic compounds such as dibenzothiophenes, various biomarker compounds, total petroleum hydrocarbons, and others. The presence and estimated concentrations of the target compounds were assembled in an Excel spreadsheet.
- Stage 1: Exploration. This stage usually starts with data preparation which may involve cleaning data, data transformations, selecting subsets of records (samples) and - in case of data sets with large numbers of variables (i.e., compounds) - performing some preliminary feature selection operations to bring the number of variables to a manageable range (depending on the statistical methods which are being considered). Then, depending on the nature of the analytic problem, this first stage of the process of data mining may involve anywhere between a simple choice of straightforward predictors for a regression model, to elaborate exploratory analyses using a wide variety of graphical and statistical methods in order to identify the most relevant variables and determine the complexity and/or the general nature of models that can be taken into account in the next stage.

  For this EPRI study, each PAH result was examined carefully for accuracy. If any data point was suspect, the raw data were reviewed and the cause of the error, is any, was corrected. Also, the data set was normalized in some instances, based on the objective of the statistical or graphical analysis being conducted.
- Stage 2 and Stage 3 activities were not conducted during this phase of the project.

Objectives

The objective of this work was to explore the existing EPRI background PAH data set for additional characteristics and compounds and trends that help delineate MGP-derived PAHs from general urban background.

2
METHODS AND MATERIALS

Source and Collection of Samples

Surface soil samples from 318 sites in 29 population centers in three states were collected and analyzed for PAHs. Site selection was conducted using a pseudo-random sampling scheme. The samples were collected from 0 to 15.2 cm (0 to 6 inches). At some locations, samples were collected from two depths, 0 - 2.54 cm and 2.54 – 15.2 cm. The samples were analyzed for wide-range hydrocarbon fingerprint by gas chromatography with flame ionization detection (GC/FID) and 43 PAHs and alkylated PAHs by GC with mass spectrometric detection (GC/MS).

The following subsections describe the sample collection approach. Also, the sample preparation and analysis methods used in the laboratory were generally the same over the several years that data were collected with some important differences. Therefore, the following subsections also evaluate the types information that might be mined from the data files based on the nature of the sample preparation and analysis methods used for each batch of samples.

Sample Collection, Preparation and Analysis Methods

Definitions

The following terms were defined specifically for the this study to minimize confusion regarding the locations from which samples were collected.

Sample Location - the actual place where an individual sample was collected. Each sample location corresponds with the sample identification number on its sample jar and on the chain-of-custody form.

Site - the property or plot of land containing one or more sample locations. For example, “Chestnut Hill Park” or “highway 1 median” was designated as the “site”.

Area - A city, town, county, or other locality that may contain multiple sites.

Urbanized Area (UA) - an area consisting of a central place and adjacent urban fringe that together have a minimum residential population of at least 50,000 people and an overall population density of 1,000 people per square mile of land area (U.S. Census Bureau, 2000).

Populated Area - an area that has a population density of 1,000 people per square mile of land area (equivalent to the UA), but with a minimum residential population of 10,000 people. The populated area designation includes smaller cities and towns that have substantial urban centers, but do not meet the U.S. Census Bureau's definition of an urbanized area.

Study Area

Because the objective of the study was to examine the distribution of PAHs in urban soils, potential sampling sites were constrained to areas with a minimum population density. The United States Census Bureau defines an urban area as having a minimum of 50,000 persons in a density of greater than 1000 persons per square mile. However, this definition would eliminate from consideration many small to medium-sized cities and towns with substantial commercial and industrial histories. Therefore, a project-specific limit, called a Populated Area, was used that included any area with greater than 10,000 persons and a density of >1000 persons/sq. mi. This included most of the smaller urban centers.

Samples were collected from four populated areas in New York State, 16 populated areas in Illinois, and 9 populated areas in Pennsylvania. The specific populated areas in Illinois to be sampled were chosen using a random selection method. Specifically, the 205 populated areas in the state (not including the City of Chicago) were sequentially numbered and then 16 were chosen using a random number generator (Microsoft Excel). Several additional populated areas were randomly selected in case one or more of the original 16 were not accessible for sampling.

In New York, four populated areas were selected in the western part of the state based in part on their position on a randomly ranked list of populated areas in the service area of the participating electric utilities, as well as judgments about the ease of gaining permission from city officials to perform the sampling and the value of performing a study in that area for the utility.

In Pennsylvania, 9 populated areas were chosen in the eastern part of the state (not including the City of Philadelphia) using the same procedure as was used in Illinois.

Site Selection

Once areas were selected for sampling, road maps and USGS 7.5 minute quad(s) for each area were used to select the actual sites or properties to be sampled. The sites within each area were also selected pseudo-randomly. A coordinate system was laid out over maps of each area and a random number generator was used to select points that fell on this coordinate system. In the field, the closest accessible site to each selected grid point was sampled. Suitable sampling sites included parks, roadway medians, utility rights-of-way, commercial properties, residential properties, parking lot buffers, or vacant lots. Sites with known or suspected releases of PAHs were not sampled. In New York, 35 sites were sampled in the first populated area, 23 sites were sampled in the second populated area, 20 sites were sampled in the third populated area, and 10 sites were sampled in the fourth populated area. For Illinois 10 sites were sampled in each of the 16 populated areas. For Pennsylvania, eight sites were sampled in each of eight populated areas and seven sites were sampled in one populated area.

Sample Locations

Two locations were sampled at each site, except for large sites, where three samples were collected. At the time of sampling, the locations were selected by the field engineer to be representative of the overall site conditions based on a visual assessment of the site. The field engineer considered the area of the site, obstructions, visible evidence of contamination, and other practical matters. Samples were not collected in proximity to known sources of PAHs, such as railroad tracks or oil storage tanks.

Sample Collection and Compositing

At every location, samples were collected from both the 0 - 2.54 cm depth interval and the 2.54 - 15 cm interval. The samples were collected with pre-cleaned stainless steel trowels by marking out a one foot square area, removing any grass or ground cover, and then carefully transferring the soil from each depth interval into pre-cleaned stainless steel bowls. Each sample was briefly mixed and any rocks, glass, wood or other debris was removed before the entire sample was transferred to a pre-cleaned soil jar. No samples were collected beneath pavement, sidewalks, or other structures. Based on this sampling scheme, four to six discrete soil samples were collected at each site.

The samples were labeled, packed on ice, and shipped to the laboratory. All of the samples were composited in the laboratory as described in the following paragraphs.

The samples from 58 of the 88 sites in New York were composited by thoroughly mixing equal weights of soil from the two samples collected from 0 to 2.54 cm at each site to generate a 0 to 2.54 composite, and similarly mixing equal weights from the two samples collected from 2.54 to 15 cm to generate a 2.54 to 15 cm composite.

For the 160 Illinois samples, two types of compositing were performed. At 32 (20%) of the sites, equal portions from the two samples collected from 0 to 2.54 cm at each site were combined, as were equal portions from the two samples collected from 2.54 to 15 cm, as was done for New York. In addition, at the other 128 (80%) of the sites, the four discrete samples were composited to create one sample for the site that represented 0 to 15 cm. Specifically, 10 grams from each of the 0 to 2.54 cm samples was mixed with 50 grams from each of the 2.54 to 15 cm samples to create a composite sample representative of the interval from 0 to 15 cm.

Finally, the remaining 30 New York samples and 71 Pennsylvania samples were composited in the same way as the 128 Illinois samples.

Sample Preparation and Analysis

Somewhat different sample preparation and analysis methods were used for the various batches of samples as described below. The nature of each preparation and analysis determines what chemical data are available in the raw data files that can be mined. For example, if sample extracts were fractionated using column chromatography prior to analysis and only the portion

containing aromatic compounds was retained and analyzed, then other potential constituents of the samples, such as saturated hydrocarbons and polar organic compounds, would have been acquired by the GC/MS. Those data could not be mined from the existing data files.

Illinois and New York Samples

Each composited soil sample from Illinois was extracted using the Soxhlet method. The non-polar and slightly polar compounds, including PAHs, were isolated from a portion of each extract using silica gel column chromatography (EPA Method 3630C). As stated above, silica gel cleanup removed compounds such as organo-phenols, -nitro, and carboxylic acids, among others.

The extracts then were analyzed by GC/MS in the full scan mode without prior cleanup or fractionation. Therefore, sufficient data is present in the instrument data files to explore for compounds not initially reported and to conduct mass spectral library searches on unidentified peaks.

Pennsylvania Samples

Each of the composited soil samples from Pennsylvania was extracted by pressurized liquid extraction (EPA Method 3545) using a Dionex ASE 300. This extraction method is generally equivalent to the Soxhlet technique. Each extract then was cleaned-up on Silica Gel and analyzed by GC/MS in the full scan mode using the same methods as were used for the New York and Illinois samples.

Quality Assurance and Quality Control

Prior to collecting samples, a quality control plan was developed that included measures, frequencies, and criteria for calibration, blanks, spikes, replicates, and other parameters. The QC criteria were based on project-specific objectives and standard U.S. EPA criteria for the applicable methods.

The reporting limits were based on the sample equivalent of the lowest linear calibration standard. The reporting limits for the parent PAHs ranged from about 4 µg/kg to about 20 µg/kg, depending on the compound and sample conditions. Detection limits were about one half of the reporting limits.

As presented in the EPRI background PAH reports (EPRI 2002, EPRI 2003a, EPRI 2004, and EPRI 2008), all data were checked for errors and quality control acceptability. Deviations from the data quality objectives were noted and appropriate sample identifiers, sample descriptions, and final data were entered into a database constructed using Microsoft® Access software.

Extraction of Alkylated PAHs and other Target Compounds from the Data Set

Using the raw electronic GC/MS data files from the previous analyses of the soil and QC samples, selected alkyl-PAHs were identified and their respective concentrations calculated using the original calibration data. The parent PAHs were present in the calibration standards used with the original analyses and were previously processed. The alkyl-PAHs were identified using their characteristic ions and pattern recognition. The concentrations of the alkyl-PAHs were calculated using the relative responses of their respective parent compounds.

In addition, several compounds of potential diagnostic value were included in the data re-processing. The concentrations were estimated using the response factor of a closely related, previously calibrated compound.

The final target analyte list is provided in Appendix A. This compound list is similar to those used by Nordtest and the US Coast guard to identify the sources of oil spills and has been demonstrated by META and EPRI to be effective at discriminating pyrogenic PAH sources in many cases (EPRI 2000a).

Potential Diagnostic Compounds Not Expected to Be in the Data Set

Ideally, in addition to the relative amounts of parent and alkylated PAHs, PAH source identification can be based on indicator chemicals found to be unique to the specific PAH sources. For example, combustion/pyrolytic products consist of parent PAH and substituted PAH, including oxygen substituted PAH, such as fluorenone, anthracenone, phenalenone, anthracenedione, naphthopyranone, and nitropyrene at detectable concentrations. Coal tar-derived and oil tar-derived PAH includes similar parent and alkylsubstituted PAH as found in combustion matter, however, the principal heteroatoms in the PAH structure are nitrogen and sulfur, rather than oxygen. Coal tar PAH assemblages include relatively high levels of the nitrogen and sulfur heterocyclics such as carbazole (and substituted carbazoles), thiophenes, benzothazole, dibenzothiophenes, benzoquinolines, and substituted acridines. However, organo-oxygen compounds, such as phenol, cresols, and substituted phenols also are present. Nitro- and keto-substituted PAHs generally are not present in coke oven coal tars (Coal Tar Data Book 1965). Byproducts of former manufactured gas plants (MGPs) utilizing oil as a feedstock contained substantially lower amounts of oxygen-containing compounds. Finally, petroleum and petroleum-derived products typically contain very low concentrations of oxygen-containing compounds, especially oxy-PAH.

Based on these differences in the chemical composition of the major PAH sources, identifying key indicator chemicals in the EPRI background PAH data set was a secondary objective. However, it was not possible to extract some of these data as discussed in the next subsections.

Cleanups

The purpose of the silica gel cleanup step was to remove the relatively large amount of natural organic matter (e.g., plant roots, decaying organic debris) present in many samples. These materials interfere with the accurate measurement of PAHs using GC/MS. However, in addition to the removal of natural organic matter, the silica gel also removed polar organic compounds such as organo-phenols, nitroPAH, and carboxylic acids, among others. Some of these compounds have potential value for identification of PAH source in urban soil. Many moderately polar organic compounds, containing nitrogen, sulfur, or oxygen atoms, also were partially or wholly retained by the silica gel.

Data files for a few samples were searched for common organic acids and none were found. Therefore, the data set was not mined for polar organic compounds.

Detection limits

The detection limts for PAHs in the original data set were approximately 10 to 25 µg/kg. The concentrations of the 16 priority pollutant PAHs ranged from not detectable to somewhat greater than 1000 µg/kg. Because many of the potential diagnostic compounds discussed above are at levels substantially less than the parent PAHs in many PAHs sources, they were not detectable by the methods used for the background PAH study. Therefore, the data mining efforts focused on compounds that are substantial constituents of common petrogenic and pyrogenic substances.

As discussed in Section 1 of this report, the GC/MS instrument responds to a greater or lesser intensity based on the structure of the particular compound being measured. This had an effect on the detectability of certain classes of compounds in the raw data files. For example, the GC/MS response factors for petroleum biomarker compounds are about 5 time less than for PAHs. Therefore, the detection limits for biomarkers were about 5 times greater; specifically, about 25 to 50 µg/kg in the soil samples. Because of this limitation of the data set, petroleum biomarkers were not detected in many sample with relatively low concentrations of PAHs.

Data Analysis

All statistical and graphical analyses were conducted using the commercial software, Microsoft Excel, Microsoft Powerpoint, and Statistica, Version 7.1, from StatSoft, Inc. (Tulsa OK).

3
RESULTS AND DISCUSSION

Site and Area Land Use Types

A total of 318 sites were sampled in New York, Illinois, and Pennsylvania, including a variety of different types of sites. Based on observations made in the field, a subjective classification of the land use was made for each site sampled as part of this study. The site use classifications included recreational (30%), rights of way (23%), municipal (24%), utility (11%), open land (7%), residential (3%), conservation (1%), industrial (< 1%), and commercial (< 1%). Recreational sites included ball fields and recreational parks. Sites where a municipal right of way existed, such as the areas near streets, were designated as rights of way even if they formed part of a residential or commercial lawn. The municipal designation was used for public areas such as police stations, fire stations, and town buildings, including schools. Conservation sites included areas such as forest and nature preserves. Sites were considered utility sites either if they were owned by a utility or if there was a utility right of way in the area sampled. The residential sites were home lots or apartment complexes. Sites with businesses on them were designated as commercial for retail business, or as industrial for manufacturing or similar operations. The open land designation was used for land with no specific purpose or structures.

All of the sites sampled were within the legal boundaries of the target populated area. For example, no open land or agricultural sites were sampled if they were outside the boundary of the particular municipality.

In addition to site use classifications, a subjective characterization of the surrounding area use was made at each of 285 sites. The resulting area use classifications included heavy residential (50%), commercial (21%), light industrial (11%), light residential (10%), agricultural (2%), heavy industrial (1%), and rural (5%). Areas were considered heavy residential if houses were present at a density of at least one house per acre. Otherwise, if the houses were sparse, the area was assigned the light residential designation. Similarly, areas with a high density of non-retail businesses were considered heavy industrial, particularly if manufacturing was present. On the other hand, areas with just one or a few non-retail businesses were considered light industrial. Commercial was used to describe areas with retail businesses. For sites bordering farmland, the area was considered agricultural. Lastly, areas with very few businesses or residences, and without farmland, were designated as rural.

Few samples were collected in areas of concentrated heavy industry, such as coke plants, steel mills, power plants, foundries, and manufacturing where the levels of PAHs in shallow soil are expected to be generally higher than in other urban areas. Also, no samples were collected on sites with obvious releases of petroleum or tar-derived substances.

Summary of Physical Sample Characteristics

Consistent with the intent of the sampling plan, the samples contained various amounts of silt, sand, clay and gravel. For the 417 individual samples classified, the TOC levels ranged from 0.2% to 19% with a median TOC of 2.7%.

Sample Results for PAHs

As stated in the Methods section, the samples were collected from 0 to 15.2 cm (0 to 6 inches). However, at some locations, samples were collected from two depths, 0 - 2.54 cm (0 to 1 inch) and 2.54 to 15.2 cm (1 to 6 inches). Further, at some locations, the 0 - 2.54 cm and 2.54 to 15.2 cm samples were analyzed separately.

A working data set was generated for the statistical and graphical analyses presented in this report. The data set was generated in two steps. First, one half the sample-specific detection limit was substituted for non-detects. Then, arithmetic composites were generated where the 0 to 2.54 cm and 2.54 to 15.2 cm samples were analyzed separately. This was done by generating weighted averages based on the sampling intervals, which are representative of the 0 to 15.2 cm interval at those locations. Finally, for some of the statistical and graphical analyses, if the measured or calculated concentration of an analyte in a sample was below sample-specific detection limit, then that result was given a null value and omitted from any further data processing.

Table 3-1 summarizes the PAH results for all 318 samples.

Summary Statistics for PAH Results from All Sites

Table 3-1 lists the summary description statistics for all 318 samples. Included in the statistics for each compound are the mean, median, upper and lower quartiles, and 5th and 95th percentiles. All statistics reported in Table 3-1 were performed on the raw data, assuming that the concentrations of each variable (PAH analyte) are normally distributed. As shown in Table 3-1, benzo(a)pyrene concentrations ranged from ND (1.9 µg/kg is ½ the detection limit) to 7,920 µg/kg with a median concentration of 130 µg/kg. The upper quartile concentration was 410 µg/kg, while the upper 95th percentile concentration was 2,220 µg/kg.

These concentrations appear to be consistent with literature-reported levels for anthropogenic background in small to medium-sized residential, commercial, and light industrial areas based on parent PAHs, but notably lower than background observed in more highly populated cities and commercial/industrial areas (Bradley 1994, MA DEP 2002). The concentrations of alkylated PAHs and other selected PAHs in Table 3-1 have rarely been reported in the literature. This is the first comprehensive report of the distribution of those compounds that we are aware of.

Table 3-1
Summary Statistics for All Samples in the EPRI Background PAH Study

Variable	Descriptive Statistics (All EPRI Background PAH Data 0-6 no dups): All Concentrations are in ug/kg dry weight												
	Valid N	Mean	Confidence -95%	Confidence +95%	Median	Minimum	Maximum	Lower Quartile	Upper Quartile	Percentile 5%	Percentile 95%	Std.Dev.	Skewness
Naphthalene	318	71	42	100	16	3.5	3008	6.0	41	3.9	214	264	9.0
2-Methylnaphthalene	318	68	38	98	13	1.9	4294	4.7	43	3.8	265	272	12.5
1-Methylnaphthalene	318	33	21	46	8	0.2	1440	3.0	21	0.7	127	114	8.6
C2 - Naphthalene	318	58	36	81	13	0.7	2503	3.8	40	1.9	206	201	8.6
C3- Naphthalene	318	51	34	68	13	1.2	1718	3.2	40	1.8	177	154	8.0
C4- Naphthalene	318	26	18	34	5	0.3	779	2.0	20	1.8	108	73	7.4
Acenaphthylene	318	42	28	57	7	1.7	1706	2.2	31	1.9	281	129	8.3
Acenaphthene	318	39	22	56	6	0.4	2149	2.1	21	1.9	161	152	10.6
Dibenzofuran	318	44	26	62	9	1.3	1927	4.2	27	3.6	154	162	8.5
Fluorene	318	47	27	67	6	0.4	2420	2.2	27	1.9	209	183	9.4
C1 - Fluorene	318	21	15	27	4	0.5	549	2.0	16	1.8	87	56	7.0
C2 - Fluorene	318	24	17	30	2	1.7	575	2.0	17	1.8	108	60	5.7
C3 - Fluorene	317	26	16	36	2	0.7	887	2.0	5	1.8	112	90	6.2
Phenanthrene	318	579	391	767	120	1.8	16870	39.3	407	5.4	2451	1702	7.3
Anthracene	318	131	88	175	21	1.4	4470	6.4	97	2.0	583	394	7.2
C1 - Phenanthrene/Anthracene	318	254	185	323	64	0.9	6238	19.4	211	2.7	1155	628	6.4
C2 - Phenanthrene/Anthracene	318	162	124	200	51	0.9	3341	14.0	156	2.0	770	344	5.4
C3 - Phenanthrene/Anthracene	318	77	61	93	28	1.7	1294	4.9	78	1.9	344	145	4.4
C4 - Phenanthrene/Anthracene	318	28	23	34	6	0.0	368	2.0	31	1.8	123	52	3.5
Dibenzothiophene	318	31	20	42	6	0.3	1108	2.2	21	1.3	120	102	8.0
C1 - Dibenzothiophene	318	21	17	26	6	0.1	341	2.0	20	0.7	95	42	4.3
C2 - Dibenzothiophene	318	27	21	33	7	1.5	375	2.1	27	1.8	125	51	3.8
C3 - Dibenzothiophene	318	18	14	22	3	0.0	244	2.0	16	1.8	84	34	3.5
C4 - Dibenzothiophene	247	4	3	5	2	0.0	63	1.9	2	1.8	15	9	4.8
Benzo(b)naphtho(2,1-d)thiophene	283	54	41	66	16	1.7	808	4.9	43	2.0	254	107	3.8
Fluoranthene	318	908	680	1136	226	1.9	19310	76.1	724	8.4	3750	2068	5.3
Pyrene	318	736	555	917	194	1.9	15356	64.4	572	7.3	2941	1641	5.4
C1 - Fluoranthene/Pyrene	318	343	265	420	102	1.3	5358	30.0	295	2.1	1596	702	4.3
C2 - Fluoranthene/Pyrene	318	190	147	233	65	1.7	3663	15.6	185	2.0	824	386	4.9
C3 - Fluoranthene/Pyrene	318	84	66	102	30	1.7	1570	6.0	90	1.9	355	165	5.0
Benz[a]anthracene	318	453	342	563	105	1.8	9159	34.6	375	4.6	2206	1000	4.9
Chrysene	318	525	407	643	146	1.9	8638	49.2	428	6.4	2282	1068	4.4
C1 - Benz(a)anthracene/Chrysene	318	354	252	456	59	1.0	7637	18.2	203	2.0	2209	925	4.5
C2 - Benz(a)anthracene/Chrysene	318	193	135	251	37	1.7	4921	5.4	117	1.9	1042	525	5.2
C3 - Benz(a)anthracene/Chrysene	318	124	83	165	13	0.5	2964	2.1	64	1.8	632	371	5.2
C4 - Benz(a)anthracene/Chrysene	318	60	38	81	2	1.7	1926	2.0	30	1.8	308	193	6.1
Benzo[b]fluoranthene	318	445	348	542	109	1.9	7007	40.3	378	5.3	2176	879	3.8
Benzo[k]fluoranthene	318	378	298	459	113	1.8	7043	38.3	364	5.4	1637	729	4.6
Benzo(e)pyrene	318	263	206	321	78	4.3	3856	24.9	217	9.7	1309	521	3.9
Benzo[a]pyrene	318	476	374	578	128	1.9	7925	47.3	410	5.3	2216	923	4.0
Perylene	317	96	75	118	25	2.8	1522	10.3	79	8.8	481	196	4.0
Indeno[1,2,3-cd]pyrene	318	337	262	412	89	4.3	6030	28.7	293	9.7	1704	678	4.1
Dibenz[a,h]anthracene	318	125	94	155	29	1.8	3008	10.7	105	8.7	568	276	5.5
Benzo[g,h,i]perylene	318	384	301	468	113	4.3	7000	33.9	355	9.8	1745	757	4.2
Benzo(b/c)fluorenes	318	34	25	44	7	0.1	908	4.1	28	1.6	192	85	5.7
2-Methylpyrene	318	49	39	59	15	0.3	683	5.0	44	1.5	224	92	3.9
4-Methylpyrene	318	45	35	56	14	0.2	813	4.5	38	1.4	187	98	5.1
1-Methylpyrene	318	34	26	41	10	0.2	585	3.1	29	1.2	141	70	4.9
total PAHs	318	8374	6487	10261	2423	125	144995	763	7783	170	36831	17103	4.5

As discussed in other EPRI reports, the differences between the mean concentrations and the median concentrations strongly suggest that the data are not normally distributed. This observation was explored further using normal probability plots and frequency histograms. Figure 3-1 and Figure 3-2 show normal probability plots for phenanthrene and total PAHs (sum of 43 PAHs and alkylPAH ranges) that clearly indicate the lognormal distribution of the data for those analytes. In constructing normal probability plots, first the data are rank ordered. Then, from these ranks, Z values (i.e., standardized values of the normal distribution) are computed based on the assumption that the data come from a normal distribution These Z values are plotted on the y-axis in the plot. If the actual PAH concentrations (plotted on the x-axis) are normally distributed, then all values should fall onto a straight line in the plot. If the values are not normally distributed, they will deviate from the line.

Similar plots were developed for the other analytes with similar results, but are not included in this report.

The results showed that PAHs in urban surface soils are log-normally distributed. The concentrations of total PAHs ranged from 125 µg/kg to 145,000 µg/kg with an average concentration of 8,370 µg/kg and a median concentration of 2,420 µg/kg. The log-normal distribution is further supported by Skewness values substantially greater than 0.

Summary Statistics for PAH Results by Site and Area Uses

Benzo(a) pyrene was used to illustrate the distribution of PAHs by site and area use. Table 3-2 and Table 3-3 show the results. The data do not show a clear trend by site use (Table 3-2). However, the data suggest that PAH concentrations are higher in areas that are industrial or commercial regardless of the land use of the individual site sampled. It is important to note, however, that the number of samples analyzed in several site use and area use categories is very small, and that the data may not be representative of those areas generally. For example, only one sample out of 318 was collected on an actual commercial site, whereas 61 samples were collected in commercial areas. The data suggest that the one commercial site result is not representative of commercial sites generally.

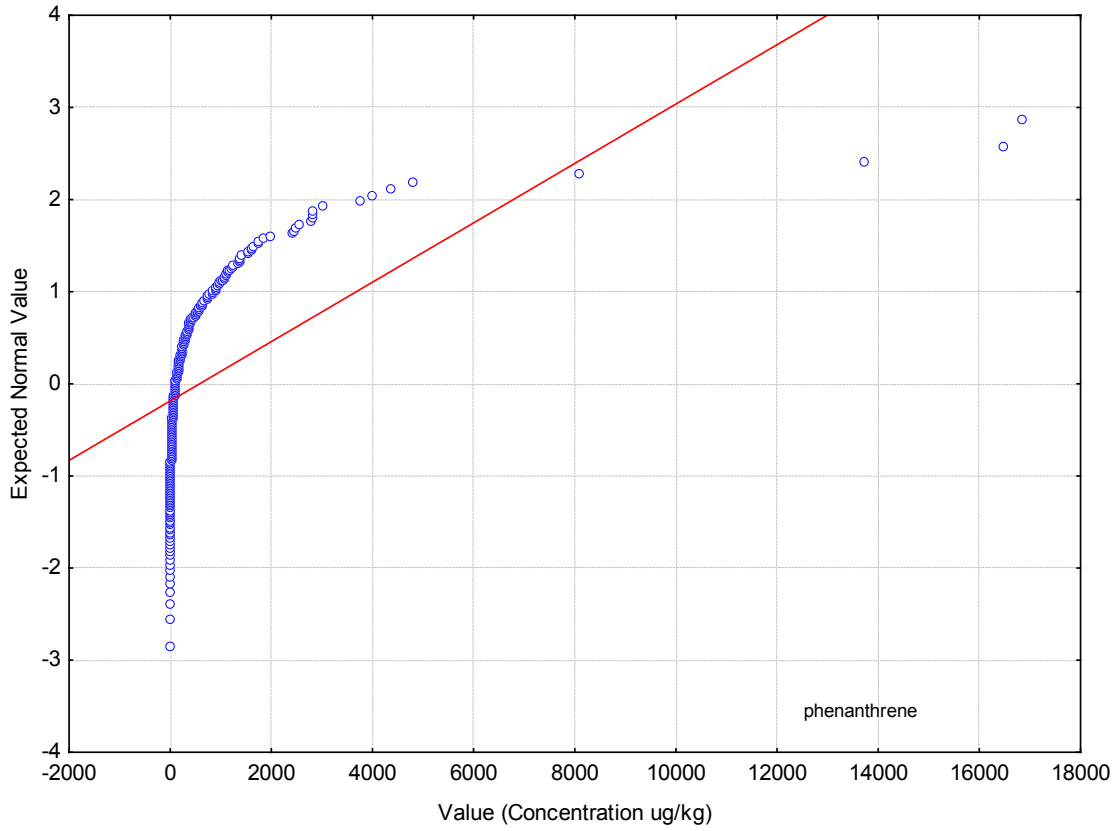


Figure 3-1
Normal Probability Plot for Phenanthrene Concentrations in Surface Soil (N=318)

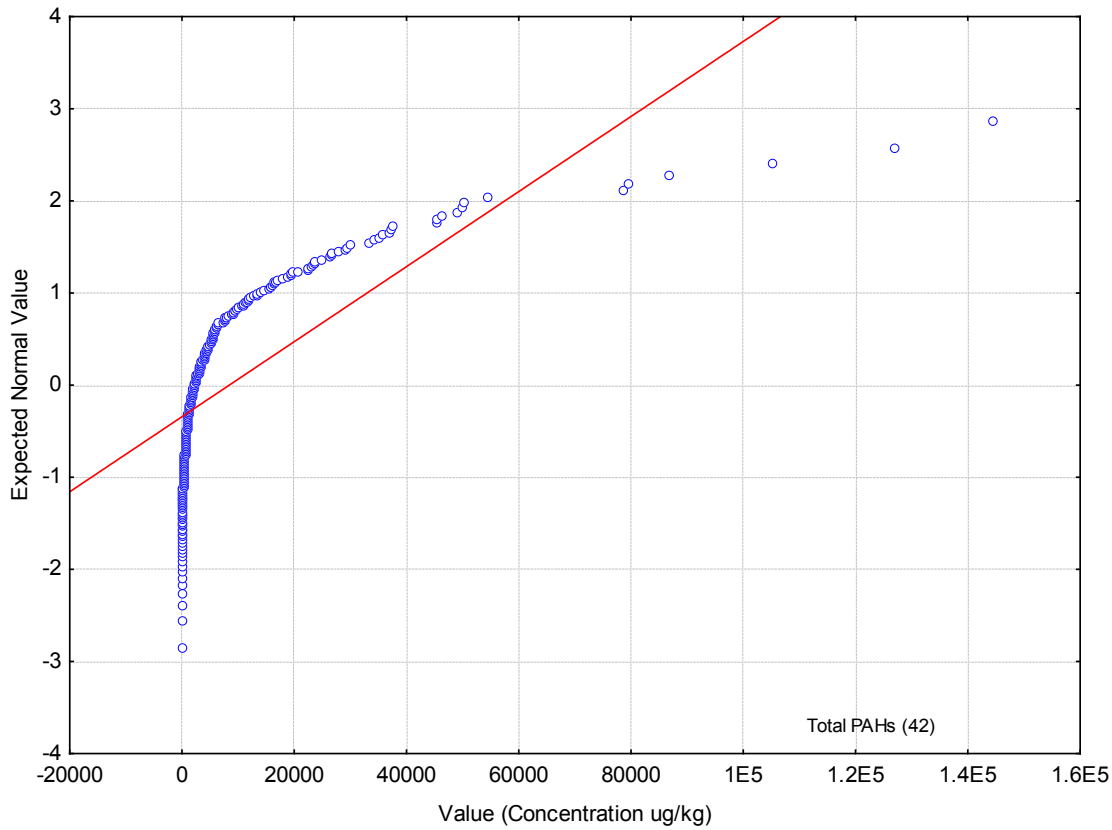


Figure 3-2
Normal Probability Plot for Total PAH Concentrations in Surface Soil (N=318)

Table 3-2
Summary Statistics by Site Uses for Benzo(a)pyrene

	N	Median (µg/kg)	Range (µg/kg)
Rights of Way	74	254	2.1 – 4,740
Recreational	96	98	2.6 – 7,920
Municipal	76	112	2.0 – 11,600
Utility	35	120	2.0 – 2,190
Open Land	22	243	2.8 – 2,960
Residential	9	108	30 – 2,150
Conservation	4	120	6.7 – 342
Industrial	2	90	70 – 110
Commercial	1	1.9	1.9 – 1.9

N = 318; results based on 0 – 15.2 cm interval and ND set to ½ DL.

Table 3-3
Summary Statistics by Area Uses for Benzo(a)pyrene

	N	Median (µg/kg)	Range (µg/kg)
Heavy Residential	142	123	2.0 – 7,920
Commercial	61	153	1.9 – 3,360
Light Industrial	32	138	2.7 – 4,740
Light Residential	27	71	2.0 – 2,220
Rural	13	30	4.9 – 1,360
Agricultural	6	68	3.3 - 135
Heavy Industrial	4	682	267 – 2,190

N = 285; results based on 0 to 15.2 cm interval; ND set to ½ DL

These data provide a PAH concentration distribution which can be described as typical urban background. PAHs found in soil within this distribution should not be assumed to be from the local MGP without further study. These data also provide a robust description of the composition of typical urban background, including such features as the relative amounts of petroleum-derived and combustion derived PAHs and the ranges of ratios of certain diagnostic PAHs and other compounds as discussed below. The data can be used, along with site-specific data, to determine whether PAHs in urban surface soil originated with local MGP operations or not. The details of the application of these techniques to MGP source attribution has been described in several journal articles and numerous scientific presentations and will not be repeated in this study.

PAH Sources and Data Trends

Total Extractable Hydrocarbons

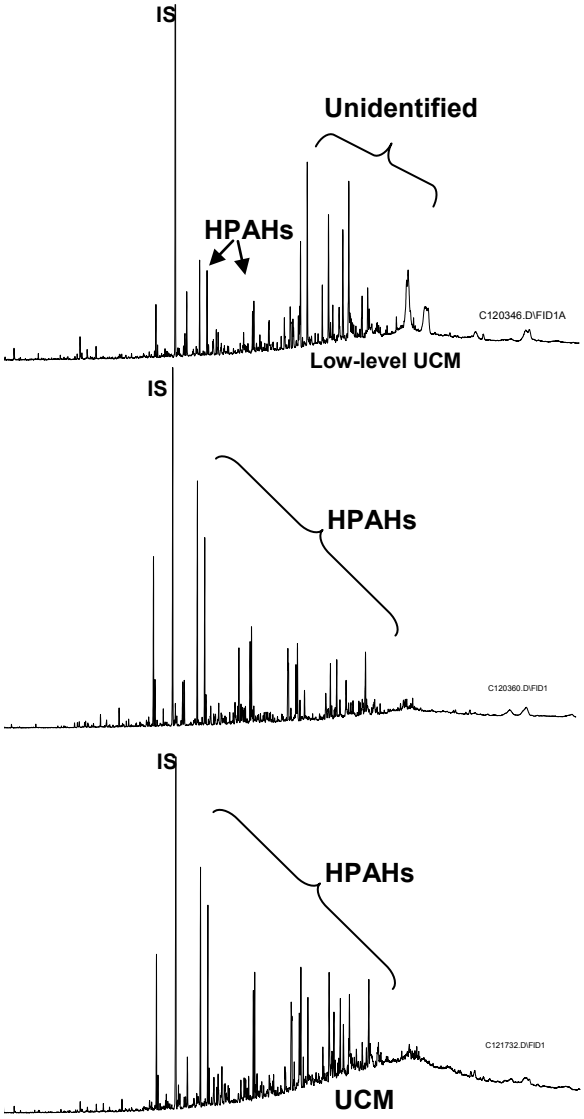
GC/FID fingerprints (chromatograms) were generated for the soil samples. The GC/FID fingerprint of a soil extract analyzed by GC/FID provides qualitative information on the amount and composition of hydrocarbons in the sample. Basic information on the whether the hydrocarbons are petrogenic, pyrogenic, or a mixture; whether they are relative low boiling or high boiling; and other characteristics are discernible.

A wide range of patterns were observed in the fingerprints from petrogenic to pyrogenic. Figure 3-5 illustrates the typical urban signatures. Most samples showed a pattern consisting of a mixture of a high molecular weight unresolved complex mixture (UCM) of hydrocarbons, indicative of the accumulation of highly weathered petroleum-derived substances and/or natural organic matter, relatively low levels of pyrogenic PAHs indicative of combustion-derived sources, and some unidentified peaks likely derived from natural plant matter.

Most samples contained high molecular weight PAHs rather than low molecular weight e.g., naphthalene, PAHs. As seen in Table 3-4, the ratio of low molecular weight PAHs (LPAHs: 2- and 3-ring compounds) to high molecular weight PAHs (HPAH: 4-, 5-, and 6-ring compounds) ranged from not calculable (no detected LPAHs) to 5.6 times as much LPAHs as HPAHs. However, the mean and median values indicate that the majority of samples have LPAH/HPAH ratios of about 0.2 or so. Figure 3-6 shows that the majority of samples had LPAH/HPAH ratios less than 0.5; only a few samples had LPAH/HPAH ratios greater than 1. In fact 93.9% of all samples had LPAH/HPAH ratios less than 0.4.

This general pattern of low LPAH/HPAH ratios and a late eluting UCM is not characteristic of MGP-derived tarry materials, nor of coke oven tars, creosote, road tar, pitches and related substances. Most of those materials contain substantial concentrations of 2- and 3-ring PAHs and little or no UCM. However, in some cases, severe weathering of tars and tar-like materials (TLM) can give compositions similar to urban background, as low molecular weight compounds degrade or preferentially wash away (Figure 3-4).

Further, mixtures of coal tar pitch and some petrogenic materials can consist of HPAHs and late-eluting UCMs principally. More source-specific information is needed to distinguish weathered TLM samples and mixtures from urban background as discussed later in this report.



IS – internal standard
UCM – unresolved complex mixture
HPAH – high molecular weight PAHs

Figure 3-3
GC/FID Chromatograms Showing Typical Urban Background Signatures

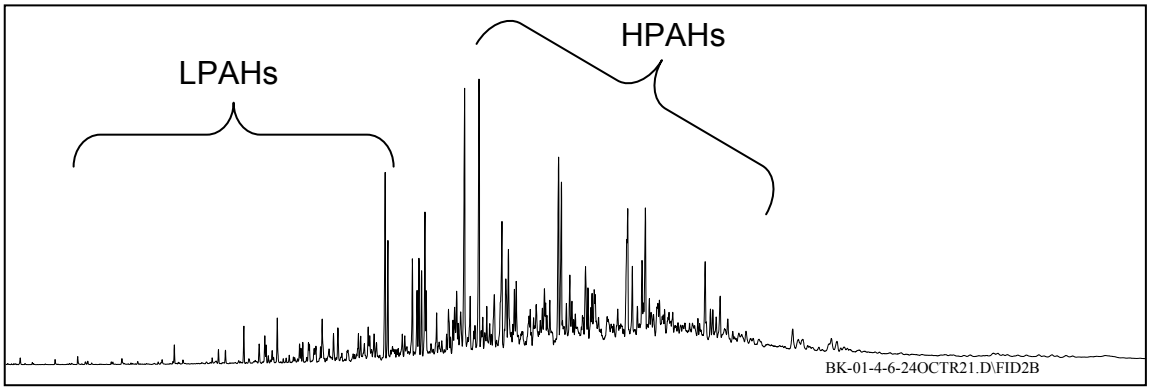


Figure 3-4
Severely Weathered TLM-Impacted Soil Sample from a Former MGP Site Showing Substantial Loss of LPAHs

Table 3-4
Descriptive Statistics for TOC and Total PAH Characteristics in Urban Background Surface Soil

Variable	Descriptive Statistics (All EPRI ratios 0-6 no dups): Concentrations in ug/kg												
	Valid N	Mean	Confidence -95%	Confidence +95%	Median	Minimum	Maximum	Lower Quartile	Upper Quartile	Percentile 5%	Percentile 95%	Std.Dev.	Skewness
Total Organic Carbon	318	32974	30369	35580	27725	2463	188000	18700	40900	9250	68083	23617	2.63
Total PP PAHs	316	5800	4345	7254	1533	15	124309	517	4730	64	21396	13144	5.60
Total PAHs	316	8536	6544	10527	2533	15	152088	750	7790	77	37104	17995	5.00
Ratio TPPPAH/TPAH	316	0.69	0.68	0.71	0.71	0.016	1.0	0.63	0.78	0.434	0.93	0.15	-0.83
LPAH/HPAH	314	0.22	0.18	0.26	0.15	0.000	5.6	0.12	0.21	0.081	0.49	0.37	11.24
Total alkylPAHs (incl DBTs)	309	2253	1746	2760	652	0.3	34659	184	1969	9.658	8841	4529	4.02

Samples with non-detects were set to null and not included in the statistic

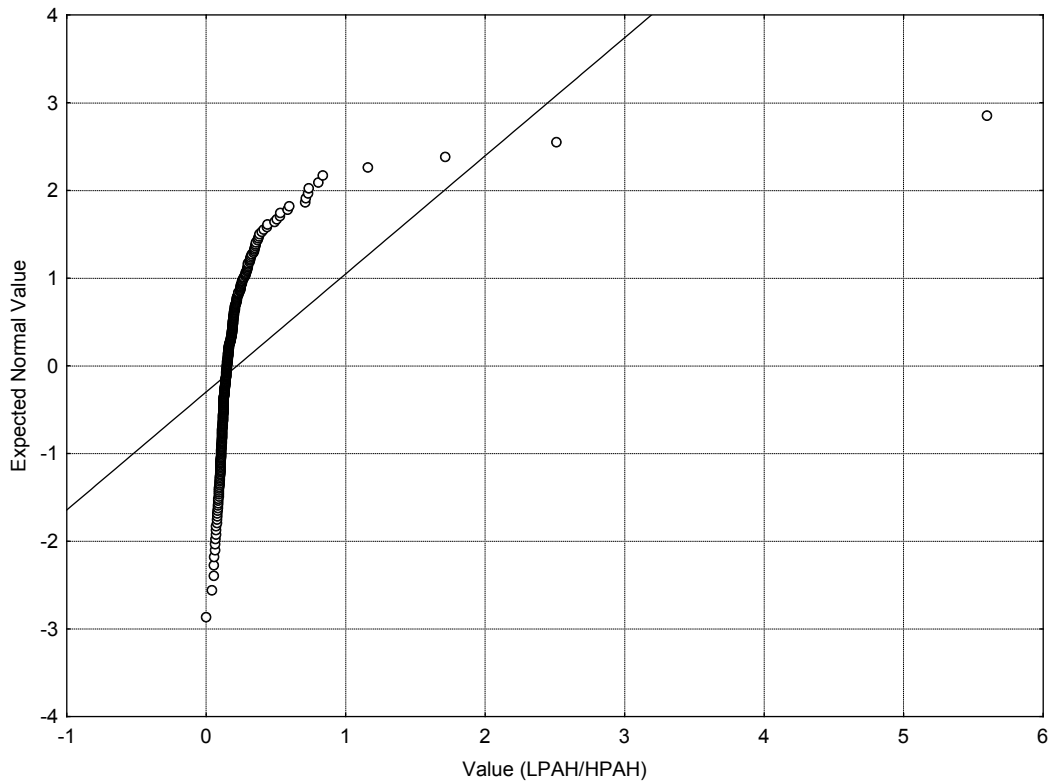


Figure 3-5
Normal Probability Plot for the Ratio LPAH/HPAH

In contrast to the distribution of LPAHs/HPAHs, the distribution of the ratio of total priority pollutant PAHs (TPPPAH) to total PAHs (TPAH) was nearly normally distributed (Figure 3-7). The average value was 0.69. A ratio near 70% is consistent with those observed for pyrogenic substances, such as MGP tars, coal tars, and creosote reported previously by EPRI (EPRI 2000a) and well above those for common petroleum products that tend to be less than 25%. This suggests that the majority of PAHs in urban background is pyrogenically-derived, either from combustion processes or coal tar-derived materials. Therefore, one focus of any study of the sources of PAHs in urban background should be determining whether the PAHs are combustion-derived or tar-derived.

Several other trends are noted in Table 3-4. For example, while total organic carbon (TOC) trends weakly with TPAHs, it is not a good surrogate for PAH impacts.

The data also show that urban background generally has less than 21,000 µg/kg TPPPAHs (95th percentile), or less than about 5 µg/kg TPPPAHs (upper quartile) when extreme values are omitted. Similarly, urban background generally has less than 37,000 µg/kg TPAHs (95th percentile) and less than 7.8 µg/kg TPAHs (upper quartile). The use of the non-parametric measures, rather than the parametric 95% confidence about the mean, is more appropriate for this data set because of the log-normal nature of the distributions. This upper bound is fuzzy, but

provides guidance for determining when to look for other more site-specific sources of PAHs at a particular site.

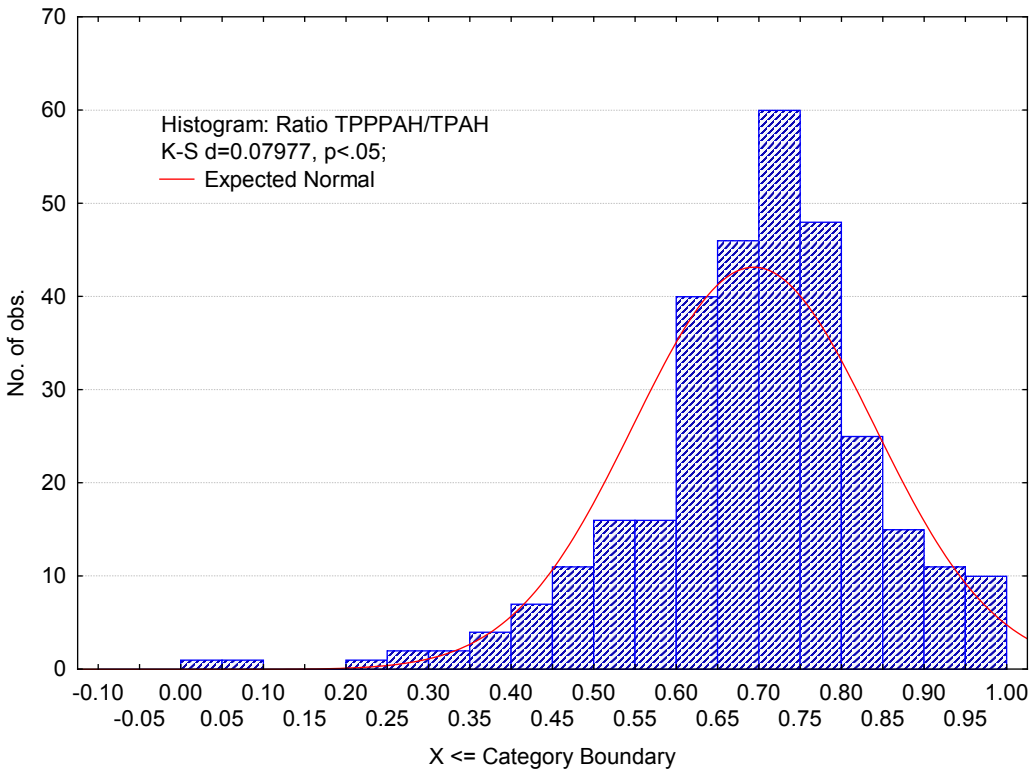


Figure 3-6
Frequency Histogram for the Ratio of Total Priority Pollutant PAHs to Total PAHs(43)

Petroleum Biomarkers

Biomarkers are complex compounds found in petroleum and coal that show little structural change from the parent compounds that were components of the organisms from which the petroleum or coal formed (Wang 2007). Different petrogenic materials have different amounts and patterns of these biomarkers making them useful compounds for environmental forensic study. Some potentially useful classes of biomarkers include isoalkanes, cyclic alkanes, terpanes, steranes, and sesquiterpanes.

Of equal importance, biomarkers are present at very low or non-detectable levels in coal tars and many former MGP TLMs (EPRI 2000a, EPRI 2003b). Also, some former MGP TLMs, derived from petroleum feedstocks, contain biomarkers that are carryover from the feedstock. Consequently, the identification of biomarkers in a soil sample suggests a petrogenic imput, at least partially.

Finally, combustion-derived PAHs originate from a variety of sources and are produced under a wide range of conditions. They are exhausted to the atmosphere where a portion precipitates out

under dry and wet conditions. At the same time, incompletely combusted fuel oil and coal are emitted to the atmosphere and also eventually precipitate out. Therefore, soil impacted primarily by combustion-derived PAHs can have both pyrogenic and petrogenic hydrocarbons.

Approximately 50% of the instrument runs for the background PAH samples were qualitatively reviewed for the presence of normal alkanes, isoprenoid alkanes (pristane and phytane), alkylated cyclohexanes, sesquiterpanes, triterpanes, and steranes. As discussed in Section 2, biomarkers could not be detected in some samples because of analysis limitations even though they likely were present. Regardless of this limitation, biomarkers were detected in many samples.

Given the uncertainties of the identification and quantitation of the petroleum biomarkers from the background PAH data set, they could not be used to estimate the relative amounts of petroleum-derived PAHs in the samples. Further, no patterns emerged that would help to assign combustion-derived PAHs to sources such as diesel exhaust or coal furnaces. However, biomarker analysis can be valuable when trying to identify MGP TLM in off-site samples because many MGP residuals contain no detectable biomarkers. Thus their presence in soil samples frequently indicates PAH inputs from non-MGP sources.

PAH Ratios

As discussed in Section 1 and earlier in this section, PAHs in urban background originate from atmospheric deposition of particulates from a number of sources, such as industrial emissions, residential and commercial heating systems, vehicles, and fires. This PAH input has occurred over many decades and is balanced to some degree by PAH degradation and transport to deeper soil zones. At some sites, this PAH deposition occurs on top of PAH-containing fill that was used at the site. As a result, the PAHs at most sites represent the combined residues from mixed sources making source attribution difficult.

One approach to the problem of MGP source apportionment might be to define urban background as the accumulation in soil of PAHs from a number of unidentified sources over long periods of time. In contrast, MGP impacts most often originate from filling activities with MGP byproducts or from migration of NAPL, contaminated soil, or water to off-site locations. A third type of impact could be defined as PAHs with a characteristic signature from other identified or likely sources such as petroleum spills, pavement sealers (USGS 2003), foundry sand, and construction debris.

By categorizing the sources into a few groups, PAH attribution is simplified into identifying the general characteristics of those groups. For example, it is clear from Table 3-5 that based on median values, the ratios of LPAHs/HPAHs are very different among the three groups. Further, the weathering of TLM or petroleum products would have to be severe before the LPAH/HPAH ratios were similar. Also, while the ratio of TPPPAH/TPAH is similar for urban background and TLM, both are different from petroleum products. These measures, along with total PAH concentrations and GC/FID fingerprints discussed above provide initial discriminating characteristics

Table 3-5
Comparison of Median and Upper and Lower Quartiles of Characteristic Ratios for Urban Background Soil Samples to Tars and Petroleum Products

	Tar and Tar-Like Materials[1]	Crude Oil and Common Petroleum Products[2]	Urban Background Soil Samples[3]
LPAH/HPAH	2.71 (2.41-4.53)	8.45 (2.76-16.9)	0.15 (0.12-0.21)
TPPPAH/TPAH	0.64 (0.55-0.70)	0.11 (0.07-0.19)	0.71 (0.63-0.78)

1 six tar samples from MGP and other sources
2 12 samples of crude oil, No. 2 fuel, and No. 6 fuel oil
3 soil samples from this EPRI study

A common approach for identifying sources of PAHs involves the use of ratios of individual compounds. Numerous papers have examined the application of PAH ratios, several focused on former MGP sites (EPRI 2000a, Yunker 2002, Costa 2005, Lima 2005, Yan 2005, Morillo 2007).

As discussed in Section 1, there are substantial differences in the relative amounts of parent and alkylated PAHs in petrogenic verses pyrogenic substances. In contrast, the differences within pyrogenic substances vary much less. The relative abundance of individual PAHs in pyrogenic substances has been shown to vary with such factors as temperature of formation, nature of the feedstock, and nature of the process (e.g., retorts v. engines v. furnaces v. uncontrolled burning). Of particular importance at former MGP sites, is the trend of increasing amounts of non-alternate PAHs (aromatic systems with at least one five member ring such as fluoranthene, benzofluoranthenes, fluorene, benzofluorenes, and indenopyrenes) relative to alternate PAHs (aromatic systems with only 6-membered rings, such as phenanthene, pyrene, etc.) (Dominguez 1996, Zander 1987). These differences allow the use of PAH ratios, such as fluoranthene/pyrene for discriminating PAH sources.

Several potentially diagnostic PAH ratios were calculated and are listed in along with their basic descriptive statistics (Table 3-6). In contrast to the overall summary statistics reported in Table 3-1, where ½ the detection limit was inserted for non-detects, the statistics in do not include samples with non-detect values. This was done to minimize bias and variance in the PAH ratio distributions, since previous studies suggest that the ratios of many PAH sources differ by only a small amount.

Fluoranthene/pyrene

Table 3-6 lists several basic statistics of the fluoranthene/pyrene ratio distribution in the background PAH dataset. The mean and median of the fluoranthene/pyrene ratio for the 316 samples with valid results were 1.184 and 1.200, respectively. Precision around the mean and median was relatively tight: the range of plus or minus one standard deviation about the mean gave a range of 1.04 to 1.32. The ratio is close to normally distributed, but with a slight skew to lower values due to the presence of some samples with abundant petrogenic matter.

Table 3-6
Descriptive Statistics for Selected PAH Ratios in Urban Background Surface Soil

Variable	Descriptive Statistics (All EPRI ratios 0-6 no dups): Concentrations in ug/kg												
	Valid N	Mean	Confidence -95%	Confidence +95%	Median	Minimum	Maximum	Lower Quartile	Upper Quartile	Percentile 5%	Percentile 95%	Std.Dev.	Skewness
Fluoranthene/Pyrene	316	1.184	1.168	1.199	1.200	0.284	1.7	1.140	1.260	0.915	1.36	0.141	-1.73
Dibenzofuran/Fluorene	191	2.081	1.436	2.726	0.828	0.096	35.4	0.564	1.485	0.412	9.31	4.521	5.52
Retene/C1-Pyrene	201	0.166	0.128	0.203	0.117	0.002	3.2	0.039	0.228	0.010	0.34	0.267	7.95
Benzofluorenes/Methylpyrenes	263	0.477	0.415	0.538	0.384	0.011	5.6	0.270	0.561	0.125	0.97	0.508	6.86
Benzo(e)pyrene/Benzo(a)pyrene	281	0.560	0.542	0.579	0.531	0.106	1.9	0.474	0.625	0.414	0.78	0.158	3.53
C2D/C2PA	200	0.174	0.164	0.184	0.160	0.023	0.4	0.126	0.225	0.067	0.30	0.070	0.63
C3D/C3PA	148	1.078	0.337	1.819	0.223	0.010	35.4	0.162	0.324	0.063	0.66	4.561	5.89
Benz(a)anthracene/Chrysene	313	0.765	0.751	0.778	0.766	0.230	1.1	0.689	0.851	0.582	0.95	0.122	-0.34
Chrysene/C2-Chrysene	247	6.062	5.336	6.787	4.851	0.014	44.6	2.788	7.617	0.520	14.8	5.786	3.12
Phenanthrene/Anthracene	277	5.961	5.662	6.260	5.497	1.433	23.0	4.349	6.781	3.033	10.6	2.526	2.01
Benzo(b)fluoranthene/Benzo(k)lfuoranthene	315	1.053	1.033	1.074	1.046	0.555	2.3	0.955	1.129	0.761	1.35	0.186	1.34
Indeno(123-cd)pyrene/Benzo(ghi)perylene	284	0.849	0.833	0.865	0.857	0.166	1.1	0.796	0.939	0.630	1.02	0.137	-1.76
Perylene/Benzo(a)pyrene	234	0.184	0.179	0.188	0.182	0.049	0.3	0.166	0.203	0.136	0.23	0.034	-0.35
Retene/Benzo(a)pyrene	202	0.133	0.104	0.163	0.107	0.001	2.6	0.029	0.191	0.005	0.25	0.211	8.52
Fluoranthene + Pyrene/HPAH	316	0.326	0.320	0.332	0.327	0.110	0.5	0.296	0.358	0.235	0.42	0.055	0.02
Alkylchrysenes/Alkylchrysenes +Benz(a)anthracene+Chrysene	290	0.317	0.294	0.341	0.261	0.030	0.99	0.190	0.350	0.108	0.791	0.203	1.41

Samples with non-detects were set to null and not included in the statistic

As discussed above, analyses of urban surface soil suggest that combustion-derived pyrogenic PAHs account for the majority of PAHs found. However, many other PAH sources are possible and some sites are expected to have PAHs from urban fill, from petroleum, and from other sources. For example, the common use of creosote treated timbers for landscaping and coal tar-based pavement sealers for driveways and parking lots have increased the tar-like character of soil at some locations. It was expected in this study that PAH ratios in background data sets would be characterized by a wide range, but with most ratios falling within a fairly narrow distribution from diffuse combustion sources. That is what is observed in Figure 3-8.

This range is well above typically fluoranthene/pyrene ratios for petroleum products (typically less than 0.4, and above those observed for former MGP TLMs from CWG plants (Figure 3-9) (EPRI 2000a, Brown 2006). Conversely, urban background fluoranthene/pyrene ratios are slightly less than observed for coke oven tars and creosote. For example, four samples of coke oven tar and creosote analyzed by META Environmental, Inc. had a mean value of 1.33 (EPRI 2000a). Finally, the fluoranthene/pyrene ratio is not a good discriminator of urban background and former coal carbonization (CC) MGP residues. Specifically, for 10 samples of MGP TLM suspected of being CC-derived, the mean fluoranthene/pyrene ratio was 1.086 with a standard deviation of 0.186. This range falls at the lower end of the urban background distribution, but overlaps the distribution substantially.

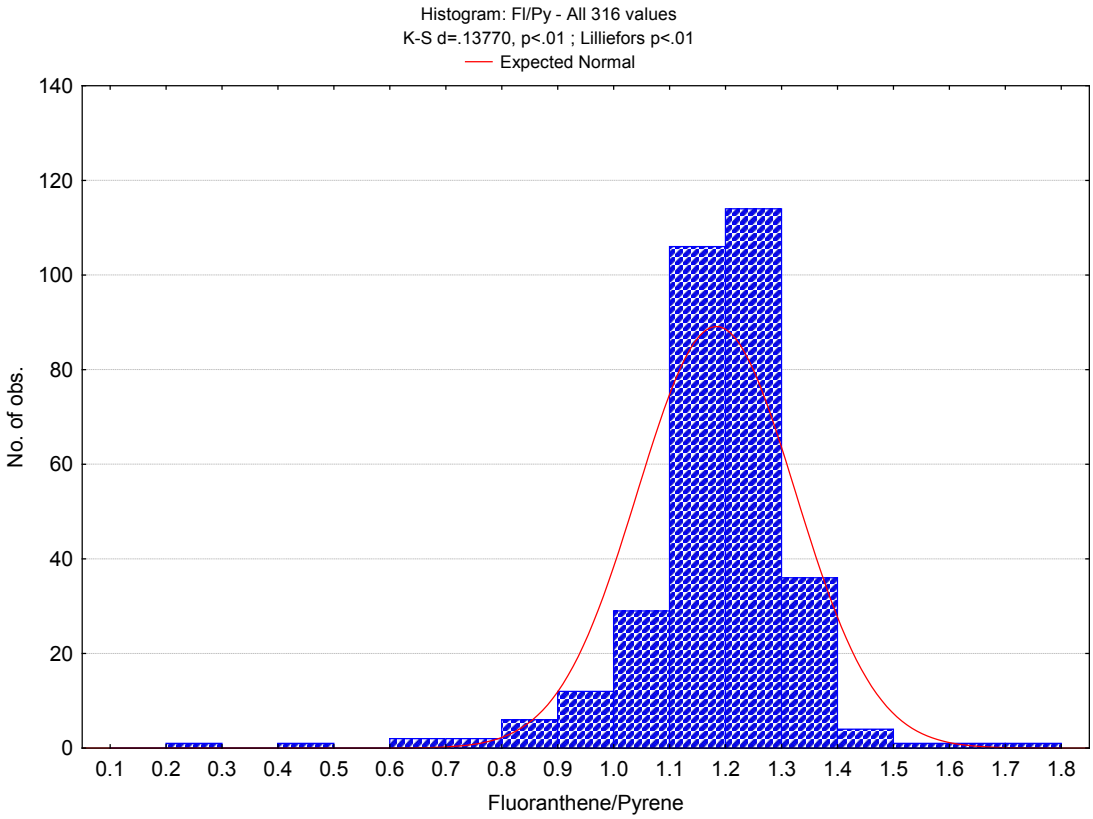


Figure 3-7
Distribution of Fluoranthene/Pyrene Ratios in Urban Background Soil

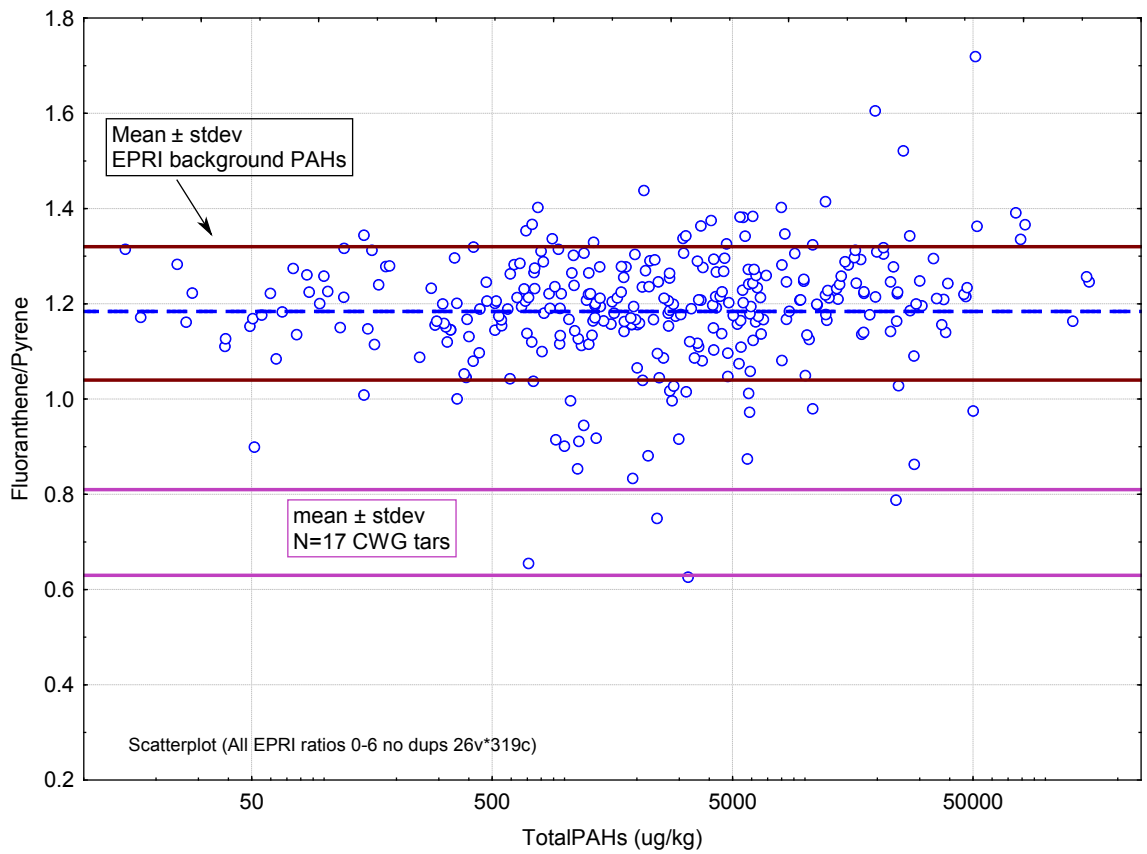


Figure 3-8
Comparison of Fluoranthene/Pyrene Ratios in Urban Background to Tar-Like Materials from Former CWG MGP Sites

For many comparisons of urban background samples and former CC MGP residuals, the discrimination is easily accomplished using the LPAH/HPAH ratio discussed above; TLM from MGP sites have substantially higher LPAH/HPAH ratios than urban background. However, for severely weathered MGP residuals, as seen in Figure 3-10, the LPAH content has been removed by environmental processes. In those cases, other characteristics of the two sources are needed.

Benz(a)anthracene/Chrysene

Benz(a)anthracene and chrysene are two alternate PAHs that vary from source to source. The benz(a)anthracene/chrysene ratio in urban background is nearly normally distributed about a mean of 0.765 (Table 3-6). When plotted against total PAH concentrations, it appears that the benz(a)anthracene/chrysene ratio increases with increasing total PAH concentration (Figure 3-11). However the trend is weak. Figure 3-11 also shows that samples of TLM from former MGP sites and from coke ovens have higher ratios than urban background.

When benz(a)anthracene/chrysene ratios are combined with fluoranthene/pyrene ratios in a double ratio plot, the discrimination between TLM and urban background becomes stronger as seen in Figure 3-12.

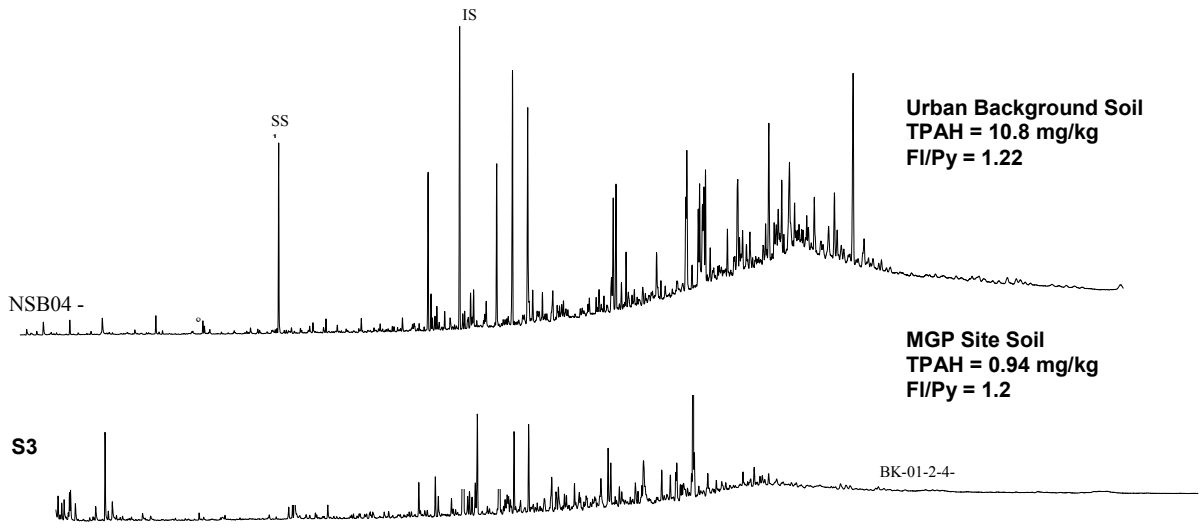


Figure 3-9
Comparison of Urban Background Soil and Severely Weathered MGP Residuals in Soil at a Former MGP Site.

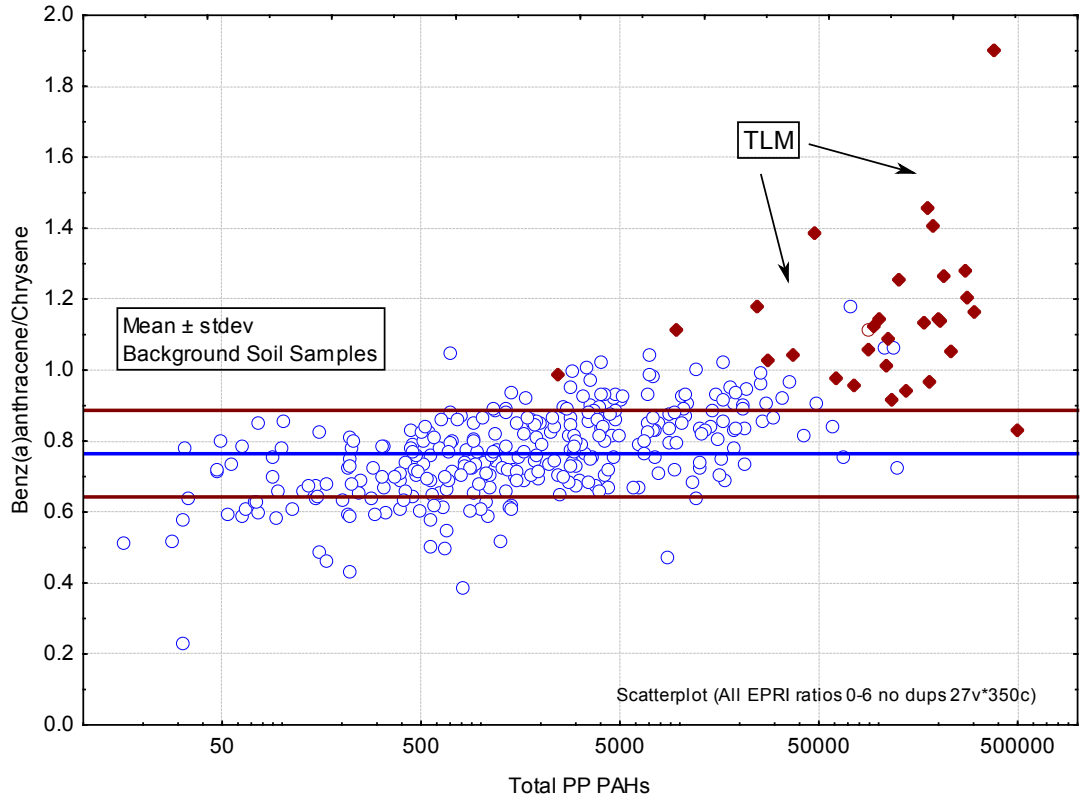


Figure 3-10
Benz(a)anthracene/Chysene Ratios of Urban Background Soil Samples and TLM

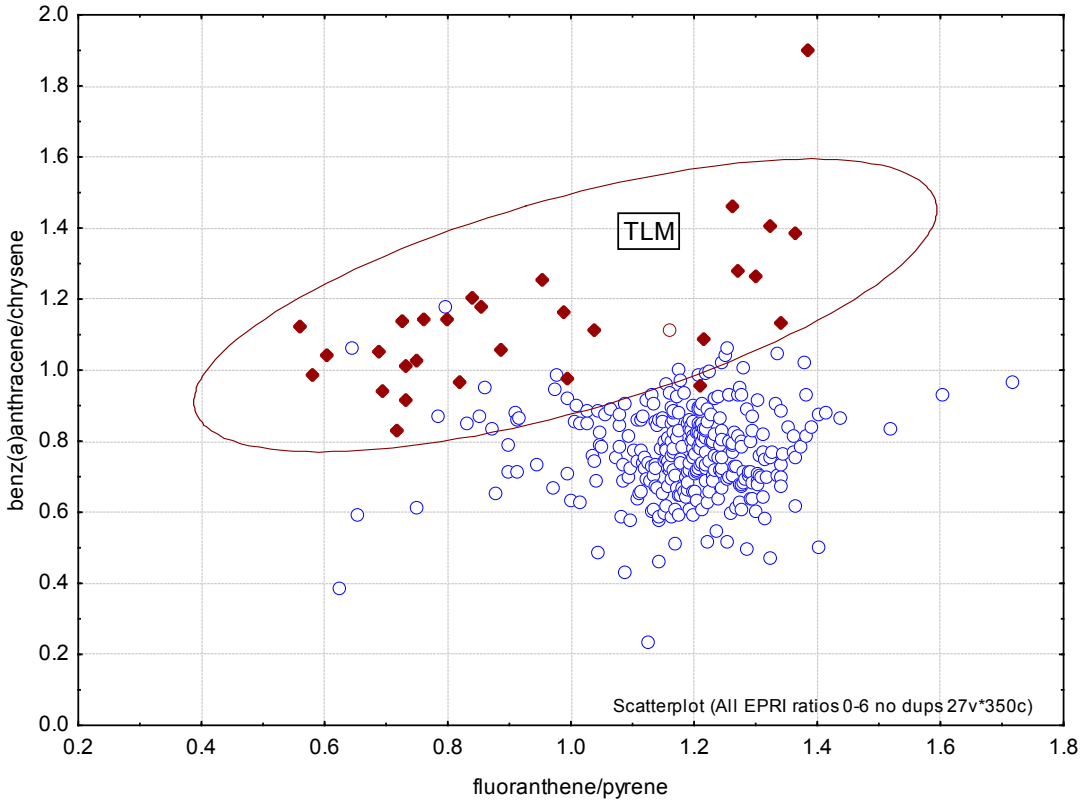


Figure 3-11
Double Ratio Plot of Benz(a)anthracene/Chrysene verse Fluoranthene/Pyrene

Benzofluorenes/Methylpyrenes

Previous research by EPRI has shown that benzo(b)fluorene and benzo(c)fluorene are non-alternate PAHs whose formation appears to have a strong temperature dependence. The abundance of benzofluorenes in TLM increases from former CWG < CC < coke tars and creosote relative to closely related methylpyrenes.

In urban background, the ratio of benzofluorenes/methylpyrenes (BF/MP) averaged 0.477 with a standard deviation of 0.508. The precision of the ratio was less than other ratios examined, with the variability increasing at both the low and high end of the concentration ranges. Comparison of benzofluorene/methylpyrene (BF/MP) ratios in urban soil to those of several TLMs in this study, indicated that the BF/MP ratio has little discriminating power for MGP source attribution generally. However, in some site-specific cases, the BF/MP ratio can be helpful. For example, the BF/MP ratio in lampblack from former west coast oil gas plants appears to be consistently low relative to urban background and other pyrogenic materials (META unpublished data).

Other PAH Ratios

Several other PAH ratios were calculated and are included in Table 3-6. These ratios were chosen because they either: 1) included combinations of alternate and non-alternate PAHs, 2) included combinations of petrogenic and pyrogenic PAHs, 3) are commonly calculated and reported in environmental forensic studies, or contained PAHs shown to be weathering-resistent (Uhler 2006).

For example, the dibenzofuran/fluorene (D/F) ratio can be useful in MGP source identification studies because dibenzofuran is an oxygen-containing heterocyclic compound found in much greater abundance in MGP coal tars, coke oven tars, and creosote than in MGP CWG tars or oil gas tars. The D/F ratio is not as useful for petroleum product identification because the refining process alters the amounts of dibenzofuran and fluorene in the various products and is not predictable. Similarly, the D/F ratio is of marginal value in urban background studies because the complex fate processes that occur in the atmosphere and shallow soil alter the ratio substantially over time creating substantial scatter among the samples.

Retene (7-isopropyl-1-methylphenanthrene) is present at significant concentrations in coal, but at much lower concentrations in coal tar or petroleum products. Thus, the ratio of retene to chrysene (or other PAHs) can be used to investigate whether coal fines are present in a soil sample and to explain some of the hydrocarbon patterns observed at sites where coal was used extensively. This is potentially important in older urban areas where coal furnaces were common in both residential and commercial buildings, and where coal dust and coal ash were dumped in backyards or added to garden soil.

As seen in Figure 3-13, urban background soils generally have substantially more retene, relative to benzo(a)pyrene, than MGP or coke oven TLM (note that the y-axis is logarithmic). It is not know from the data in this study whether the elevated retene was from deposition of PAHs derived from coal combustion or other sources. However, when combined with other ratios, such as benz(a)anthracene/chrysene, ratios that include retene can be helpful to identify potential MGP impacts in surface soil.

The ratio of benzo(e)pyrene/benzo(a)pyrene (BeP/BaP) is stable in the environment and has been used to discriminate PAH sources at former MGP sites (META unpublished data). However, as seen in Figure 3-14, this ratio appears to have little discriminating power for background sources. The values of the representative MGP and other TLMs fall well within the distribution of BeP/BaP ratios observed in the urban background soil.

The ratios of alkylated dibenzothiophenes (C2D, C3D) and alkylated phenathrenes (C2PA, C3PA) are used extensively in oil spill source identification because of their presence and variability in crude oil. The alkylated dibenzothiophenes are loosely related to the organo-sulfur content of the petroleum. The ratios also vary among former MGP residuals (EPRI 2000a). However, the concentrations of these range compounds are relatively low in TLM. As seen in Figure 3-15, the alkylated dibenzothiophene-based ratios provide little general discriminating power for MGP TLM. However, on a site-specific basis or when combined with other chemicals

or ratios, the dibenzothiophenes can help separate MGP-derived from background PAHs in some cases.

The ratio of phenanthrene/anthracene (P/A) has been proposed as a reliable way of distinguishing petrogenic from pyrogenic PAHs (Yunker 2002, Lima 2005). However, as pointed out by some authors, the ratio is susceptible to alteration in the environment (Costa 2005, Uhler 2006). Also, as discussed above, much of the PAHs in urban background are pyrogenically formed, not petrogenically formed. Therefore, the source identification objective is not to distinguish petrogenic from pyrogenic PAHs, but rather, combustion-derived PAHs from MGP-derived PAHs, both being pyrogenic. Because of the degradation potential of the P/A ratio, it was not examined in this study.

The ratio of benzo(b)fluoranthene/benzo(k)fluoranthene (BBF/BKF) has been suggested as a source diagnostic ratio. However, as seen in Figure 3-16, there is little discrimination between urban background and MGP and other TLM using this ratio.

The ratio of the sum of fluoranthene and pyrene to the sum of high molecular weight PAHs (HPAHs) is a potentially useful diagnostic. This ratio is essentially the fraction of HPAHs accounted for by fluoranthene and pyrene, the two most abundant PAHs in background samples. In general, combustion-derived PAHs appear to have a lower proportion of fluoranthene and pyrene than MGP and other TLMs. Figure 3-17 clearly shows the difference in this ratio between the urban background samples and 26 samples of TLM from a number of sites.

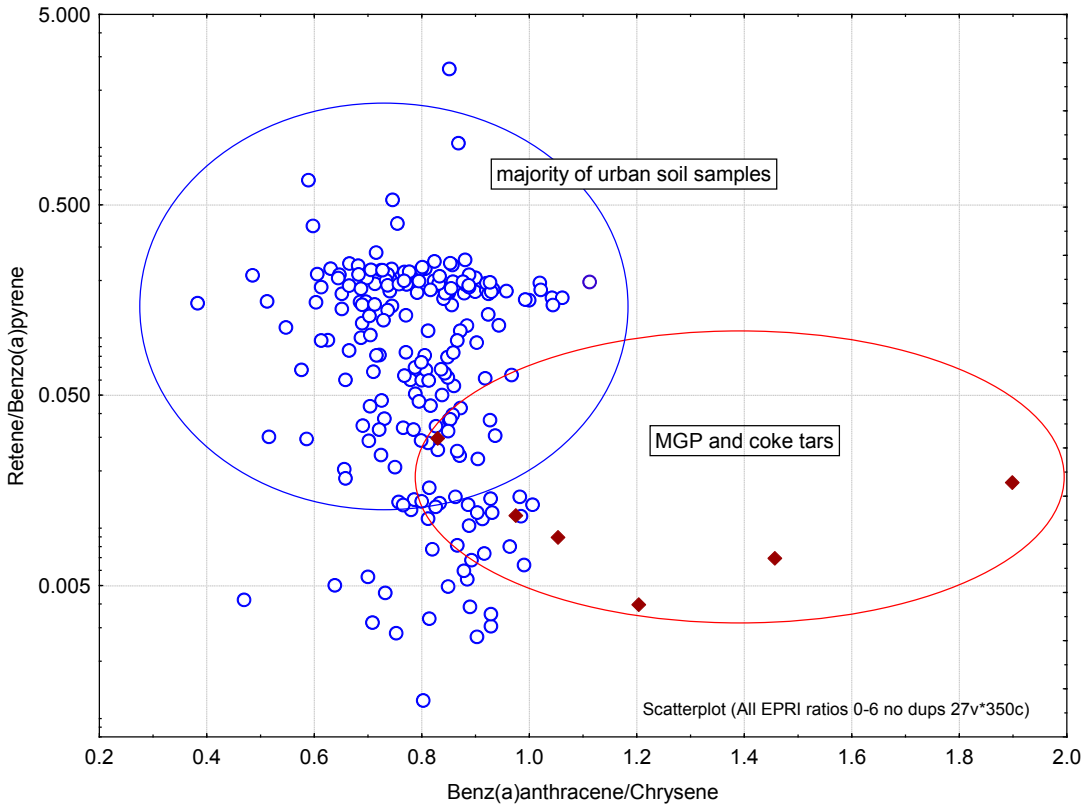


Figure 3-12
Double Ratio Plot: Retene/Benzo(a)pyrene verses Benz(a)anthracene/Chrysene for Urban Background Soil and TLM Samples

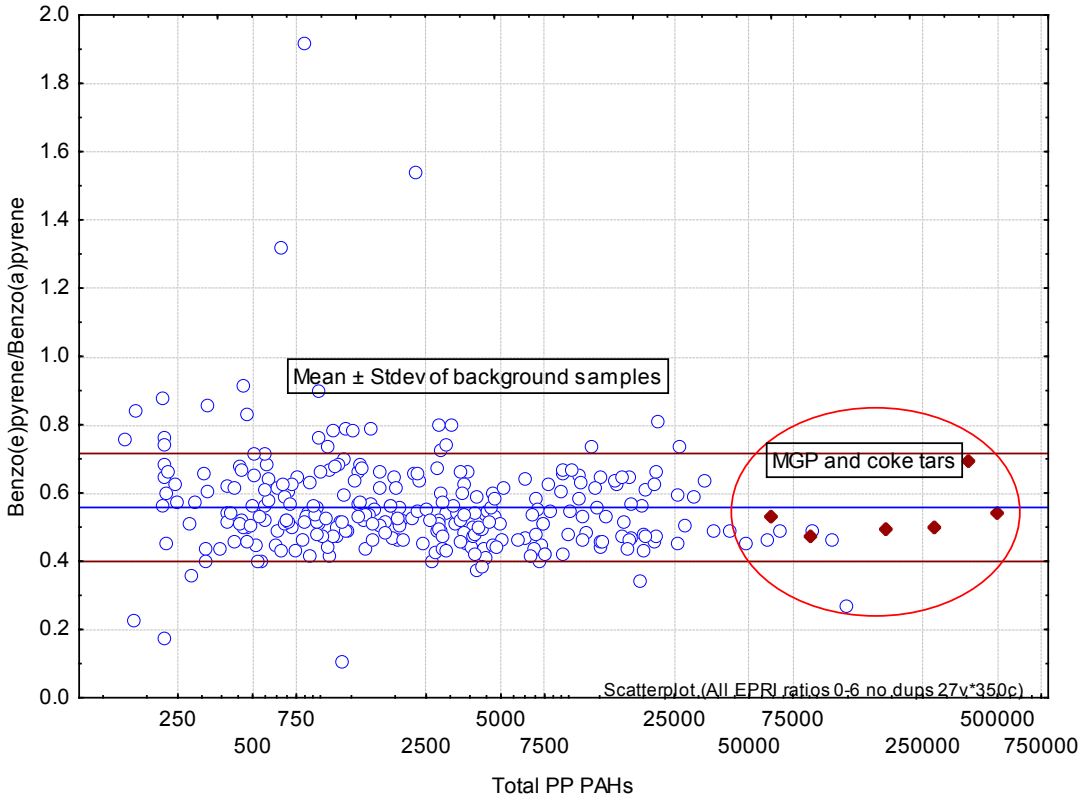


Figure 3-13
Benzo(e)pyrene/Benzo(a)pyrene Ratio in Urban Background Soil and TLM Samples

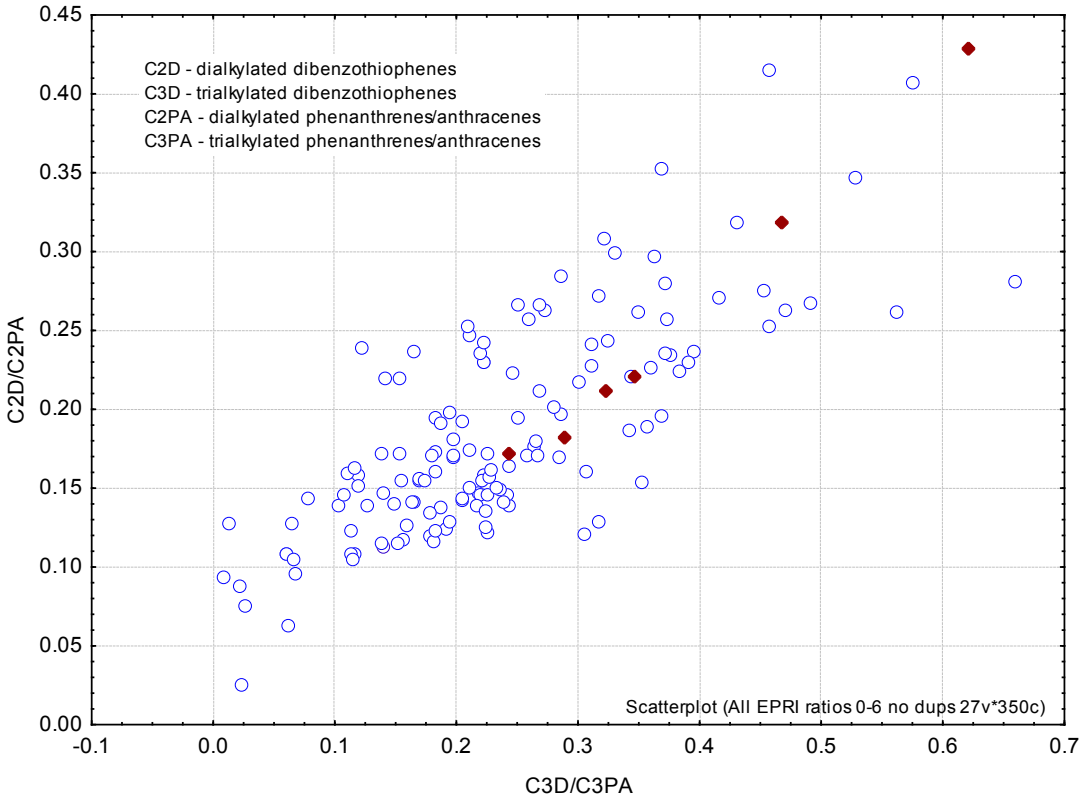


Figure 3-14
Trialkylated-dibenzothiophenes/Trialkylated-phenanthrenes verses Dialkylated-dibenzothiophenes/Dialkylated-phenanthrenes

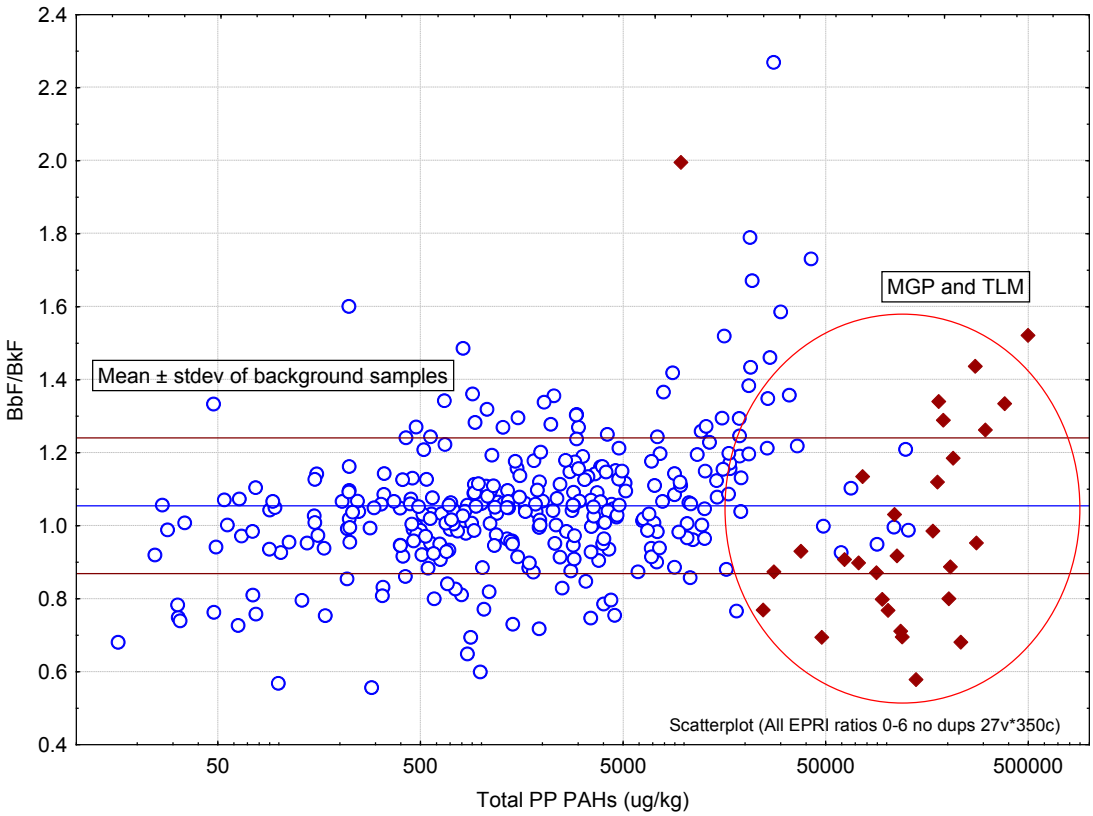


Figure 3-15
Ratio of Benzo(b)fluoranthene/Benzo(k)fluoranthene verses Total PAHs in Urban Background Soil and Selected TLM Samples

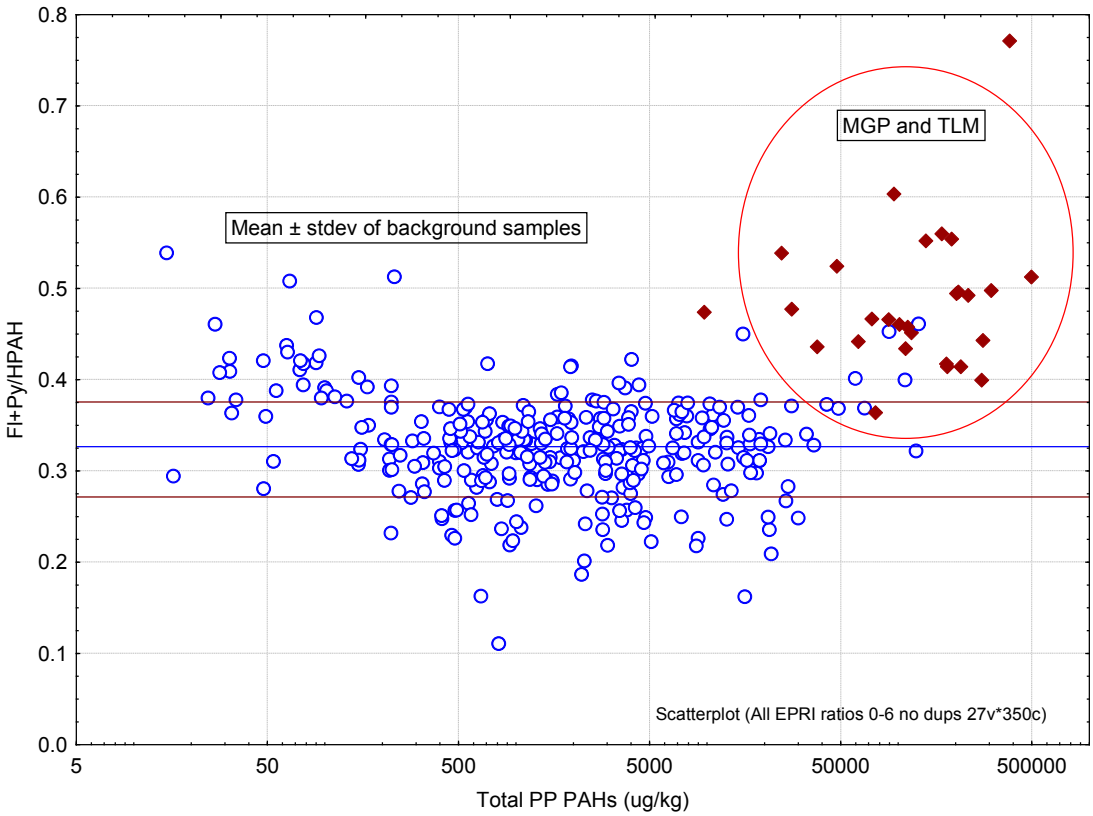


Figure 3-16
The Ratio of Fluoranthene+Pyrene/HPAHs verses Total PP PAHs in Urban Background Soils and TLM Samples

Some PAH ratios appear to show consistent differences between urban background and TLM from a number of sources. The EPRI background PAH data set includes a large number of samples collected from several areas and therefore should represent the variability in the various ratios observed in urban background generally. When extreme values are omitted, simple statistics provide a various levels of confidence that a PAH ratio value at a particular site, and therefore the PAHs in that sample, are background.

Multivariate Exploratory Analysis

A Principal Components Analysis (PCA) was conducted on the full background PAH data set as follows. First, ½ the reporting limit was substituted for non-detects. Those samples that had total PAH concentrations less than 500 µg/kg were omitted from the data table because they contained too many non-detects for valid analysis. Then, each sample was unit normalized (the sum of all compounds was one). Finally, the data were mean centered and subjected to PCA using the correlation matrix (Statistica, StatSoft, version 7.1).

As expected, PCA indicated up to eight or more PAH patterns contributing to the variance in the data. However, some of the variance was likely due to substitution for non-detects and variations in analytical data over the many batches of analyses. As discussed above, the number

of potential PAH sources are many, with small differences in PAH patterns contributing to variance. Figure 18 shows the projections of the samples onto the first two principal components. A main cluster of samples is shown that includes mixtures of HPAHs in a pyrogenic pattern and some petrogenic character. Sample points trending to the lower left contain increasing amounts of primarily pyrogenic HPAHs. Sample points trending to the upper left contained very low levels of PAHs and many non-detects. No new patterns or relationships among the compounds in the data set were descerned from the PCA.

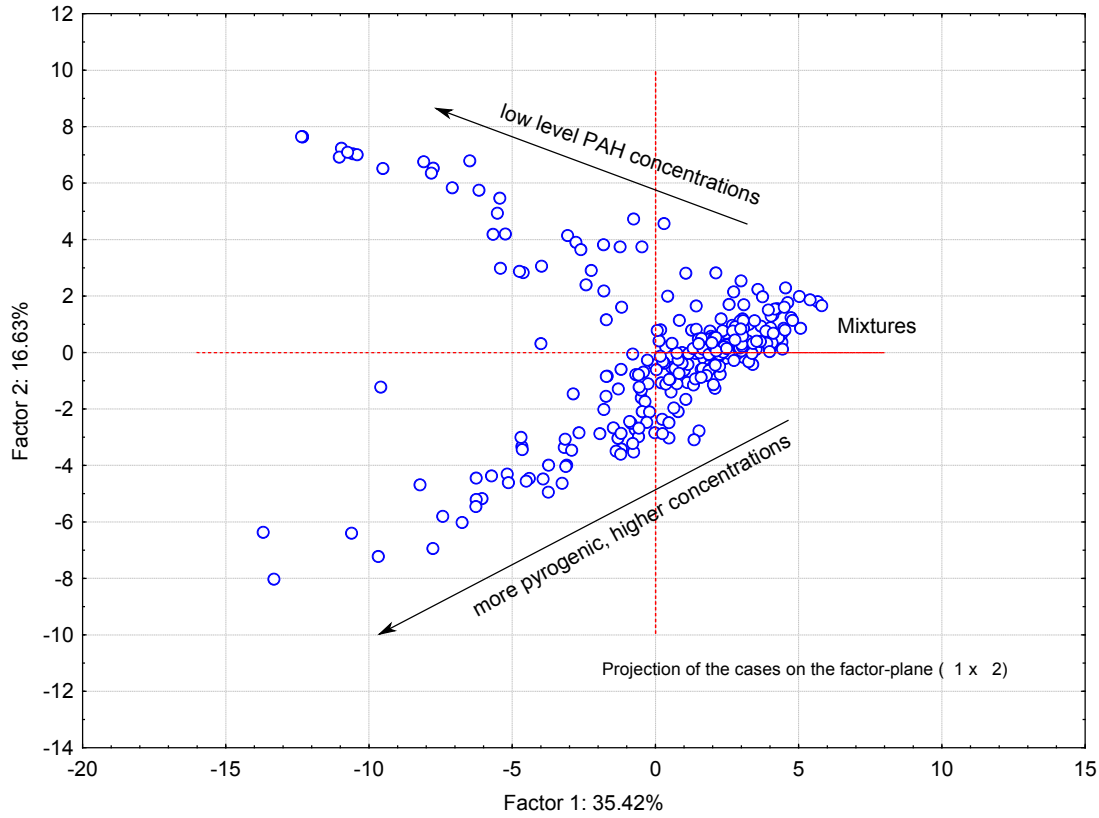


Figure 3-17
Principal Components Analysis of Normalized Background PAH Data Set

4
SUMMARY

A detailed review of the chemical analysis data for the 318 surface soil samples from 29 population centers in three states in the EPRI database was conducted. Based on that review and an examination of the recent literature, the following were concluded.

- The potential sources of PAHs in urban background is large when one considers the number of potential combustion sources, petroleum sources, coal, urban fill, and other materials. Further, deposition of PAHs from atmospheric sources has occurred over decades and is continuing to occur. Based on this conceptual model, urban background is a mixture of PAHs from potentially many sources. The EPRI study did not sample sites with known direct inputs from major sources of PAHs, e.g., fuel spills, coal tar products or wastes, industrial dumps, etc., and therefore provides a statistical picture of the accumulation of these diffuse PAH sources.
- As reported in prior EPRI reports, the concentrations of individual PAHs and total PAHs were lognormally distributed; the statistics of which are summarized in Table 3-1 for the 48 compounds and range hydrocarbons extracted from the data. The concentration distributions for the range hydrocarbons (alkylated PAHs) and selected diagnostic compounds (e.g., dibenzothiophene, retene, benzo(b/c)fluorenes) also were lognormally distributed. A comparison of the GC/FID fingerprints of high concentration samples to intermediate concentration samples shows that the EPRI data set contains a few relatively high concentration samples with a coal tar-like PAH pattern and many other samples with a combustion-like PAH pattern. The specific source(s) of the PAHs in the high concentration samples could not be identified; however, the patterns were generally consistent with coal tar products, such as weathered creosote, road tar, and pavement sealers.
- The distributions of LPAHs were more skewed than the distributions of HPAHs due to the much greater percentage of non-detect and very low concentration detection for LPAHs. Similar results were obtained for range hydrocarbons and selected diagnostic compounds, indicating that those compounds originated from the same sources as the parent PAHs in general.
- Urban background is characterized by a pattern consisting of a mixture of high molecular weight hydrocarbons, indicative of the accumulation of highly weathered petroleum-derived substances and/or natural organic matter, relatively low levels of pyrogenic PAHs indicative of combustion-derived sources, and some unidentified peaks likely derived from natural plant matter. A relative few samples contained higher concentrations of PAHs in patterns suggestive of coal tar and coal tar products. These samples may have contained contaminated fill or may have been impacted by coal tar sources such as creosote or pavement sealers.
- The ratio of LPAHs/HPAHs in urban background is low, (median = 0.15) with 95% of all samples less than 0.49. This is in contrast to crude oil, common petroleum products, and

former MGP TLM, that all have LPAH/HPAH ratios substantially greater than 0.15. The differences in LPAH/HPAH ratios among potential sources hold whether parent PAHs are used or whether the expanded list of parent and alkylated PAHs is used. Because weathering can remove the LPAHs from former MGP residuals, the LPAH/HPAH ratio is not an absolute discriminating characteristic. However, if the LPAH/HPAH ratio is greater than about 0.5, the probably is very high that some or all of the PAHs in the sample were from a non-background source.

- Some PAH ratios or combinations of PAH ratios can be used to discriminate background PAHs from MGP-derived PAHs. Some PAH ratios have little discriminating power. Of particular value, the ratio of fluoranthene/pyrene (Fl/Py) falls between about 0.5 and 0.8 for TLM from former CWG MGPs and falls between about 0.9 and 1.3 for urban background. However, the Fl/Py ratio for former CC MGP tars also falls between about 0.9 and 1.3. Therefore, other ratios are needed. For example, the ratio of benz(a)anthracene/chrysene (BAA/CHR) is greater in CC MGP TLM and coal tar and creosote from coking operations than it is urban background. Other PAH ratios with some discriminating power include retene/benzo(a)pyrene and the sum of fluoranthene and pyrene verses total HPAHs.
- It is important to note that, while some PAH ratios do not provide good discriminating power in general based on the data in this report, they may provide discriminating power on a site-specific basis where the potential MGP source is known and can be characterized more narrowly.
- PAH ratios developed for oil spill identification, including C3D/C3PA and C2D/C2PA, do not provide much power to source identification of background PAHs because much of the PAH load in urban background is combustion-derived, not petroleum-derived. However, the presence of petroleum biomarker compounds can be an indication of non-MGP PAH sources because many MGP TLMs do not contain petroleum biomarkers.
- The EPRI background PAH data can be used to assign basic statistical confidence to the various concentration ranges and PAH ratios.

This study provides MGP site owners with a high quality, robust data set to use as a benchmark or baseline when trying to determine whether PAHs found in urban soil near their sites originated with the former MGP operation. If they suspect that the PAHs originated with another operation or are part of the general urban background, then MGP site owners have several methods available with which to attribute the PAHs to those sources as discussed throughout the report.

5
REFERENCES AND BIBLIOGRAPHY

Blumer, M., "Polycyclic Aromatic Compounds in Nature," Scientific American, Vol. 243, pp. 35-45 (1976).

Boehm, P., *Polycyclic Aromatic Hydrocarbons (PAHs)*, In: Environmental Forensics Contaminant Specific Guide. Academic Press, Burlington MA, 2006.

Bradley, L.J.N., B.H. Magee, and S.L. Allen, "Background Levels of Polycyclic Aromatic Hydrocarbons (PAHs) and Selected Metals in New England Urban Soils". Journal of Soil Contamination. Vol. 3, pp. 349-361 (1994).

Brown, D.G., Gupta, L., Kim, T., Moo-Young, H.K., Coleman, A.J., "Comparative assessment of coal tars obtained from 10 former manufactured gas plant sites in the Eastern United States," Chemosphere, Vol. 65, 2006.

Costa, H.J., White, K.A., and Ruspantini, J.J., "Distinguishing PAH Background and MGP Residues in Sediments of a Freshwater Creek," Environmental Forensics, Vol. 5, 2004.

Costa, H.J. and Sauer, T.C., "Forensic Approaches and Considerations in Identifying PAH Background," Environmental Forensics, Vol. 6, 2005.

Dominguez, A., Alvarez, R., Blanco, C.G., Diez, M.A., "Chromatographic evaluation of some selected polycyclic aromatic hydrocarbons of coal tars produced under different coking conditions and pitches dermived from them," Journal of Chromatography A, Vol. 719; 1996.

Chemical and Physical Characteristics of Tar Samples from Selected Manufactured Gas Plant (MGP) Sites, EPRI, Palo Alto CA, Report TR-102184, May 1993.

Chemical Source Attribution at Former MGP Sites., EPRI, Palo Alto, CA, NYSEG, Binghamton, NY, and RG&E, Rochester, NY. 2000a. 1000728.

Literature Review of Polycyclic Aromatic Hydrocarbons, EPRI, Palo Alto, CA: 2000b. TR-114755.

Polycyclic aromatic hydrocarbons (PAHs) in surface soil, EPRI, Palo Alto, CA. 2002. Report 1005295.

Polycyclic Aromatic Hydrocarbons (PAHs) in Surface Soils in Western New York: Follow-up Report for New York State Soils, EPRI, Palo Alto, CA, New York State Electric & Gas

Corporation, Binghamton, NY, Rochester Gas and Electric Corporation, Rochester, NY, and Public Service Enterprise Group Services, LLP, Newark, NJ: 2003a. 1005296.

Identifying PAHs from Manufactured Gas Plant Sites, EPRI, Palo Alto, CA. 2003b. 1005289.

Polycyclic Aromatic Hydrocarbons (PAHs) in Surface Soils in Illinois: Background PAHs, EPRI, Palo Alto, CA, We Energies, Milwaukee, WI, and IEPA, Springfield, IL: 2004. 1011376.

Polycyclic Aromatic Hydrocarbons in Surface Soils in Eastern Pennsylvania. EPRI, Palo Alto, CA and Exelon Corporation, Chicago, IL: 2008. 1018272.

Gilbert, R. O., 1987. Statistical Methods for Environmental Pollution Monitoring. John Wiley & Sons, Inc., New York, NY.

Lima, A.L.C., Farrington, J.W., and Reddy, C.M., "Combustion-Derived Polycyclic Aromatic Hydrocarbons in the Environment – A Review," Environmental Forensics, Vol 6, 2005.

Mauro, D.M., DeClercq, P.J., Siegener, R., Coleman, A., "Survey of the distribution and sources of PAHs in urban surface soils," Land Contamination & Reclamation, Vol. 14, 2006.

Massachusetts Department of Environmental Protection (MA DEP). May 2002. Technical Update – Background Levels of Polycyclic Aromatic Hydrocarbons and Metals in Soil.

McCarthy KJ, Emsbo-Mattingly S, Stout SA, Uhler AD. Identifying Manufactured Gas Plant residues in industrial sediments. Soil Sediment & Groundwater 2000;Oct/Nov:1-3.

Morillo, E., Romero, A.S., Maqueda, C., Madrid, L., Ajmone-Marsan, F., Greman, H., Davidson, C.M., Hursthouse, A.S., Villaverde, J., "Soil Pollution by PAHs in urban soils: a comparison of three European cities," Journal of Environmental Monitoring, vol. 9, 2007.

Oil Spill Environmental Forensics Fingerprint and Source Identification. Z. Wang an S. Stout, Eds., Academic Press, Burlington, MA 2007.

Rabideau, A.J., Bronner, C., Milewski, D., Golubski, J., Weber, A.S., "Background Concentrations of Polycyclic Aromatic Hydrocarbon (PAH) Compounds in New York State Soils," Environmental Forensics, Vol. 8, 2007.

Stout SA, Uhler AD, McCarthy KJ. PAH can provide a unique forensic "fingerprint" for hydrocarbon products. Soil & Groundwater Cleanup 1998;October:1-4.

Stout, S.A., Uhler, A.D., and Emsbo-Mattingly, S., Comparative Evaluation of Background Anthropogenic Hydrocarbons in Surficial Sediments from Nine Urban Waterways, Environ. Sci. Technol., 2004, 38:2987-2994

The Coal Tar Data Book. Second Edition, The Coal Tar Research Association, Percy Lund, Humphries & Co. Ltd., London. 1965.

Uhler, A.D. and Emsbo-Mattingly, S.D., "Environmental Stability of PAH Indices in Pyrogenic Tars," Bulletin of Environmental Contamination and Toxicology, Vol. 76, 2006.

U.S. Environmental Protection Agency, 1986. *Test Methods for Evaluating Solid Waste Physical/Chemical Methods*. SW-846 Third Edition. Office of Solid Waste, U.S. EPA, Washington, D.C.

U.S. Environmental Protection Agency, 2002. Guidance for Comparing Background and Chemical Concentrations in Soil for CERCLA Sites, OSWER 9285.7-41, EPA 540-R-01-003.

USGS. *Concentrations of Polycyclic Aromatic Hydrocarbons (PAHs) and Major and Trace Elements in Simulated Rainfall Runoff From Parking Lots, Austin, Texas*, USGS Open-File Report 2004-1208. 2003.

Yan, B., Abrajano, T.A., Bopp, R.F., Chaky, D.A., Benedict, L.A., Chillrud, S.N., "Molecular Tracers of Saturated and Polycyclic Aromatic Hydrocarbon Inputs into Central Park Lake, New York City," Environmental Science & Technology, Vol. 39, 2005.

Yunker, M.B., Macdonald, R.W., Vingarzan, R., Mitchell, R.H., Goyette, D., Sylvestre, S., "PAHs in the Fraser River basin: a critical appraisal of PAH ratios as indicators of PAH source and composition," Organic Geochemistry, Vol. 33, 2002.

Zander, M., "On the composition of pitches," Fuel, Vol. 66, 1987.

A
TARGET ANALYTE LIST

Naphthalene
2-Methylnaphthalene
1-Methylnaphthalene
C1-Naphthalene
C2-Naphthalene
C3-Naphthalene
C4-Naphthalene
Acenaphthylene
Acenaphthene
Dibenzofuran
Fluorene
C1-Fluorene
C2-Fluorene
C3-Fluorene
Dibenzothiophene
C1-Dibenzothiophene
C2-Dibenzothiophene
C3-Dibenzothiophene
C4-Dibenzothiophene
Phenanthrene
Anthracene
C1-Phenanthrene/Anthracene
C2-Phenanthrene/Anthracene
C3-Phenanthrene/Anthracene
C4-Phenanthrene/Anthracene
Fluoranthene
Pyrene
C1-Fluoranthene/Pyrene
C2-Fluoranthene/Pyrene
C3-Fluoranthene/Pyrene
Benz(a)anthracene
Chrysene
C1-Benz(a)anthracene/Chrysene
C2-Benz(a)anthracene/Chrysene
C3-Benz(a)anthracene/Chrysene
C4-Benz(a)anthracene/Chrysene
Benzo(b)fluoranthene
Benzo(k)fluoranthene
Benzo(e)pyrene
Benzo(a)pyrene
Perylene
Indeno(1,2,3-cd)pyrene
Dibenz(a,h)anthracene
Benzo(g,h,i)perylene
1-Methylpyrene
2-Methylpyrene
4-Methylpyrene
Benzo(b)naphtho(2,1-d)thiophene
Benzo(b/c)fluorenes
Coronene
Retene

Export Control Restrictions

Access to and use of EPRI Intellectual Property is granted with the specific understanding and requirement that responsibility for ensuring full compliance with all applicable U.S. and foreign export laws and regulations is being undertaken by you and your company. This includes an obligation to ensure that any individual receiving access hereunder who is not a U.S. citizen or permanent U.S. resident is permitted access under applicable U.S. and foreign export laws and regulations. In the event you are uncertain whether you or your company may lawfully obtain access to this EPRI Intellectual Property, you acknowledge that it is your obligation to consult with your company's legal counsel to determine whether this access is lawful. Although EPRI may make available on a case-by-case basis an informal assessment of the applicable U.S. export classification for specific EPRI Intellectual Property, you and your company acknowledge that this assessment is solely for informational purposes and not for reliance purposes. You and your company acknowledge that it is still the obligation of you and your company to make your own assessment of the applicable U.S. export classification and ensure compliance accordingly. You and your company understand and acknowledge your obligations to make a prompt report to EPRI and the appropriate authorities regarding any access to or use of EPRI Intellectual Property hereunder that may be in violation of applicable U.S. or foreign export laws or regulations.

The Electric Power Research Institute (EPRI), with major locations in Palo Alto, California; Charlotte, North Carolina; and Knoxville, Tennessee, was established in 1973 as an independent, nonprofit center for public interest energy and environmental research. EPRI brings together members, participants, the Institute's scientists and engineers, and other leading experts to work collaboratively on solutions to the challenges of electric power. These solutions span nearly every area of electricity generation, delivery, and use, including health, safety, and environment. EPRI's members represent over 90% of the electricity generated in the United States. International participation represents nearly 15% of EPRI's total research, development, and demonstration program.

Together...Shaping the Future of Electricity

Program:

MGP Site Management

© 2008 Electric Power Research Institute (EPRI), Inc. All rights reserved. Electric Power Research Institute, EPRI, and TOGETHER...SHAPING THE FUTURE OF ELECTRICITY are registered service marks of the Electric Power Research Institute, Inc.

Printed on recycled paper in the United States of America

1015558

Electric Power Research Institute
3420 Hillview Avenue, Palo Alto, California 94304-1338 • PO Box 10412, Palo Alto, California 94303-0813 USA
800.313.3774 • 650.855.2121 • askepri@epri.com • www.epri.com