

Los Alamos National Laboratory

ENVIRONMENTAL RESTORATION

2266 General

University of California
Environmental Restoration, MS M992
Los Alamos, New Mexico 87545
505-665-4557
FAX 505-665-4747

U. S. Department of Energy
Los Alamos Area Office, MS A316
Los Alamos, New Mexico 87544
505-665-7203
FAX 505-665-4504

Date: **May 1, 1995**
Refer to: EM/ER: 95-155

Ms. Barbara Driscoll
NM/Federal Facilities Section
Environmental Protection Agency, Region 6
1445 Ross Avenue, Suite 1200
Dallas, TX 75202-2733

**SUBJECT: INTERIM GUIDANCE FOR EVALUATING POLYCYCLIC
AROMATIC HYDROCARBONS IN SOIL**

Dear Ms. Driscoll:

Enclosed for your review is an interim guidance memorandum written by the Los Alamos National Laboratory's Environmental Restoration (ER) Assessment Council entitled, "Evaluating Polycyclic Aromatic Hydrocarbons (PAHs) in Soil," dated April 12, 1995. If at all possible, we would appreciate your comments by May 5, 1995. The memorandum is at an early stage of development, but ER personnel need immediate guidance on making decisions regarding PAHs in soil. We have developed a method of comparison similar to our statistical comparisons to backgrounds for metals. When we receive your comments on the enclosed draft, we will work toward a position paper for dealing with PAH comparisons.

If you have any questions regarding this document, please call Alison Dorries at (505) 665-4791. Thank you.

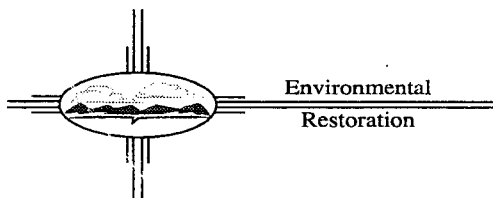
Sincerely,

Tracy Glatzmaier
Tracy Glatzmaier
Environmental Restoration Project

Sincerely,

Court Fesmire
Court Fesmire
Los Alamos Area Office

TG/CF/bp



12610

April 27, 1995

- Enclosures: (1) Draft Interim Guidance, "Evaluating Polycyclic Aromatic
Hydrocarbons (PAHs) in Soil"
(2) References

Cy (w/o enc.):

A. Dorries, TSA-11, MS K557
C. Fesmire, LAAO, MS A316
J. Jansen, EM/ER, MS M992
D. McInroy, EM/ER, MS M992
T. Taylor, LAAO, MS A316
EM/ER File, MS M992
RPF, MS M707

MAY 19 1995

Mr. Theodore Taylor
Program Manager
Environmental Restoration Program
Department of Energy
Los Alamos Area Office
Los Alamos, NM 87544

Re: *Interim Guidance for Evaluating Polycyclic Aromatic
Hydrocarbons in Soil*
Los Alamos National Laboratory (NM0890010515)

Dear Mr. Taylor:

Enclosed are comments from the Environmental Protection Agency's (EPA) review of the draft document entitled *Interim Guidance for Evaluating Polycyclic Aromatic Hydrocarbons in Soil*. Los Alamos National Laboratory (LANL) has used a study (Bradley et al. 1994) from which data was collected for polycyclic aromatic hydrocarbons (PAHs) in New England. The selection and use of this data set is not appropriate. LANL may choose to collect their own background data for PAHs; however, EPA recommends that prior to conducting the actual study, EPA review the proposed sampling locations.

Should you have any questions, please contact Barbara Driscoll at (214) 655-7441.

Sincerely,

William K. Honker, P.E.
Chief, RCRA Permits Branch

Enclosure

cc: Mr. Benito Garcia
New Mexico Environment Department
Mr. Jorg Jansen
Los Alamos National Laboratory, MS M992

6H-PN:BDRISCOLL:BD:5/17/95:J:\USER\SHARE\LANL.PAH FILE:TECH

6H-PN	6H-PN	6H-P
OWEN	NELEIGH	MIXON

General Comments
Interim Guidance for Evaluating
PAHs in Soil

1. The establishment of background data is for use in the determinations made for remedial action, not for determination of potential chemicals of concern. In that regard, the need for "Interim Guidance" (page 1) and the use of surrogate background data sets may not be necessary and the appropriate effort at this time should be to proceed with actual background sampling of polycyclic aromatic hydrocarbons (PAHs).
2. This document begins by stating that no actual background data for the LANL site have been collected. Although this document states (pg 1) that various publications exist on background concentrations of PAHs and that, ideally, the surrogate background data set should be comprised of data collected under conditions as close as possible to site-specific conditions at LANL, a data set from New England (Bradley et al. 1994) is selected for use. The selection and use of this data set does not appear to be appropriate. New England is known to have a greater likelihood of anthropogenic sources of PAHs than the southwest. For instance, many of the coal-fired electrical generating plants in northeastern U.S. have been around for 20+ years and some have not been required to meet recent air emission standards.
3. Other problems exist with this data set. Although the summary values of the original data suggest that the data set is strongly skewed to the right (most samples resulted in low background concentrations), LANL wishes to select a upper tolerance limit (UTL) as the concentration for comparison and define the UTL as the 95th percent upper confidence limit on the 99th percentile of the "estimated" distribution (NOTE: the actual data set is not provided nor are any details of the Bradley background study such as location, potential anthropogenic bias, etc.). Use of this UTL will result in specific PAH background values that are in some cases roughly an order of magnitude higher than the mean of that specific PAH of the surrogate data set. Note that the arithmetic means of the surrogate data set PAHs is much closer to the minimum values than the maximum values. In addition, use of this methodology for calculation of the UTL can result in a UTL higher than any sample analysis in the surrogate background data base (indeed, this has occurred for benzo(b)fluoranthene).

4. What is the documentation for the use of the 99% percentile as the tolerance interval? The risk assessment guidelines (RAGS) for Superfund recommend hypothesis testing with statistical significance generally determined at $P = 0.05$ for comparative testing of background samples with potential release sites (PRS).
5. For six of the PAHs (benzo(a)pyrene, benzo(a)anthracene, benzo(b)fluoranthene, benzo(k)fluoranthene, dibenzo(a,h)anthracene and Indeno(1,2,3-cd)pyrene), the action levels are lower than the proposed UTLs for background. These chemicals should remain in the risk assessment process until risk management decision-making.
6. The *Statistical Comparisons to Background, Part 1* states (pg 2, paragraph 2) that "...guidance on statistical comparisons between PRS data and naturally occurring concentrations of metals in soils provided by this document may also be applicable to background comparisons of certain anthropogenic compounds...". The very next sentence states "...it should be emphasized that the approach presented in this document should be applied to naturally-occurring metals only...". This is contradictory. Either the methodologies of this document are appropriate for use with PAHs or they are not.

DRAFT

DRAFT MEMORANDUM

Subject: Interim Guidance for Evaluating Polycyclic Aromatic Hydrocarbons in Soil

The purpose of this memorandum is to provide Environmental Restoration (ER) Project personnel with an overview of interim guidance for evaluating polycyclic aromatic hydrocarbons (PAHs) detected in soil at the Los Alamos National Laboratory (the Laboratory) potential release sites (PRSs). This guidance is intended to be finalized by mid-1995. Until that time, ER Project personnel are cautioned that evaluations completed according to this draft interim guidance may require revision by Assessments Council personnel at a later date. It should also be noted that this interim guidance does not address potential ecological concerns. Therefore, the procedures described in this document should not be implemented without the assistance of a risk assessor and statistician from the Assessments Council.

Basis for Interim Guidance

Because PAHs are present in the environment as a result of both natural and anthropogenic sources, the Assessments Council is attempting to develop a consistent strategy for distinguishing the concentrations of PAHs that can be reasonably attributed to background sources from those that are the result of past Laboratory releases. If PAH concentrations at a PRS are indistinguishable from existing background concentrations, then remedial action may not be necessary.

Information regarding background levels of PAHs in soil in the vicinity of the Laboratory is not currently available. This interim guidance presents a phased approach that will initially rely upon a surrogate background data set based on data published in the scientific literature. As data specific to the Laboratory become available, the surrogate information will be re-evaluated, and the background distributions will be adjusted as necessary. Ultimately, sufficient site-specific data that are representative of local background conditions will be collected and inclusion of surrogate data will no longer be necessary.

Surrogate Background Data Set

Various researchers have published background concentrations of PAHs in rural, agricultural, and/or urban soils collected throughout the United States and the world. Ideally, the surrogate background data set would be comprised of data collected under conditions as close as possible to site-specific conditions at the Laboratory. The Assessments Council has reviewed several existing data sets; however, most are not applicable to LANL and/or do not contain sufficient information to

be useful (e.g., analysis of benzo(a)pyrene only). At this time, the Assessments Council has adopted data from a single source (Bradley et al., 1994) to be used as the surrogate background data set. It should be noted, however, that the surrogate set is based on published summary statistics alone because Bradley et al. have not released the raw data to the Laboratory. This data set was chosen because:

- the analyses includes all individual PAHs of interest to the ER Project,
- the analytical methods and quality assurance/quality control measures are consistent with those used by the ER Project, and
- the data were collected specifically to identify background concentrations of PAHs in urban soil.

The Bradley et al. data are summarized in Table 1, and are the **only** data to be used at this time. As noted above, this data set will be amended as more information becomes available. A new memorandum will be issued by the Assessments Council each time the surrogate background data set is refined.

TABLE 1

Compound	Minimum Detect (mg/kg)	Maximum Detect (mg/kg)	Arithmetic Mean (mg/kg)	Frequency of Detection	
2-Methylnaphthalene	.017	0.64	0.151	19	62
Acenaphthene	.024	3.40	0.201	30	62
Acenaphthylene	.018	1.10	0.173	24	62
Anthracene	.029	5.70	0.351	54	62
Benzo(a)anthracene	.048	15.00	1.319	58	62
Benzo(a)pyrene	.040	13.00	1.323	57	62
Benzo(b)fluoranthene	.049	12.00	1.435	55	62
Benzo(g,h,i)perylene	.200	5.90	0.891	36	62
Benzo(k)fluoranthene	.043	25.00	1.681	59	62
Chrysene	.038	21.00	1.841	60	62
Dibenzo(a,h)anthracene	.020	2.90	0.388	32	62
Fluoranthene	.110	39.00	3.047	60	62
Fluorene	.022	3.30	0.214	35	62
Indeno(1,2,3-cd)pyrene	.093	6.00	0.987	43	62
Naphthalene	.018	0.66	0.125	35	62
Phenanthrene	.071	36.00	1.838	61	62
Pyrene	.082	11.00	2.398	61	62

Comparing PRS Data to the Surrogate Background Data Set

In early 1995, the Assessments Council issued a draft position paper on background comparisons for metals and radionuclides (*Statistical Comparisons to Background, Part I*, Los Alamos National Laboratory, Environmental Restoration Project, Assessments Council, LA-UR-95-1217). The draft background position paper describes two types of comparisons: a "hot measurement" test involving comparison of the maximum detected concentration with an estimated upper tolerance limit (UTL) of the background distribution, and a "distributional shift" test involving a comparison of the distribution of measured concentrations to the entire background distribution. At this time, the Assessments Council is recommending that only the hot measurement test be used for PAHs, because:

- UTLs can be estimated from the summary data presented in Table 1, and
- the distributional shift test requires raw data, which Bradley et al. have not agreed to release to the Laboratory.

It should be noted that an addendum to the Assessments Council's position paper on background comparisons is currently in development, and will contain several additional statistical tests that may be used to perform background comparisons. One or more of the additional tests may be applicable in cases where the raw data are not available. If so, a revised memorandum regarding PAHs will be issued by the Assessments Council.

The UTLs estimated from Bradley et al.'s summary data are presented in Table 2; the method used to calculate the UTLs is presented in Appendix A. In general, if the maximum concentration detected at a PRS falls below the UTL of the surrogate background concentration for a particular PAH, then no further evaluation of that PAH should be necessary. It is strongly recommended, however, that historical site information be considered in this decision, especially if the measured concentrations of all PAHs are below their respective UTLs. In addition, the presence of other potential chemicals of concern may also affect the decision to further evaluate PAHs, even if PAHs are present at concentrations below the UTLs.

If the maximum concentration for a PAH exceeds the corresponding UTL, then one of several possible actions may be undertaken, including, but not necessarily limited to:

- comparing the maximum detected concentration to the appropriate screening action level (SAL),

- identifying the PAH as a chemical of potential concern to be included in a baseline risk assessment,
- collecting PRS-specific background data for PAHs during Phase 2 sampling, or
- performing additional statistical analyses to confirm that the site data are different than background.

TABLE 2

COMPOUND	UTL (mg/kg)*	SAL (mg/kg)
2-Methylnaphthalene	0.64	NA [°]
Acenaphthene	3.40	4,800
Acenaphthylene	1.10	NA
Anthracene	4.29	24,000
Benzo(a)anthracene	12.40	1
Benzo(a)pyrene	12.10	0.1
Benzo(b)fluoranthene	12.20	1
Benzo(g,h,i)perylene	5.90	NA
Benzo(k)fluoranthene	19.40	1
Chrysene	19.50	96
Dibenzo(a,h)anthracene	2.90	0.1
Fluoranthene	32.50	3,200
Fluorene	3.30	3,200
Indeno(1,2,3-cd)pyrene	6.00	1
Naphthalene	0.66	3,200
Phenanthrene	24.20	NA
Pyrene	12.80	2,400

* The UTL is used except when more than 20% of the samples had non-detectable amounts of the compound of interest, then the maximum observed value is given.

[°] Not available.

The SALs for the PAHs are also presented in Table 2. Of the 17 individual compounds,

- 7 PAHs have SALs greater than the UTLs by a factor of at least 5 (acenaphthene, anthracene, chrysene, fluoranthene, fluorene, naphthalene, and pyrene),
- 4 PAHs have no SAL due to lack of toxicity information [2-methylnaphthalene, acenaphthylene, benzo(g,h,i)perylene, and phenanthrene], and

DRAFT

- 6 PAHs have SALs lower than the estimated UTLs [benzo(a)anthracene, benzo(a)pyrene, benzo(b)fluoranthene, benzo(k)fluoranthene, and dibenzo(a,h)anthracene, and indeno(1,2,3-cd)pyrene].

The relationship between UTLs and SALs will be taken into account when deciding the need for further action.

Conclusions

This memorandum provides an overview of interim guidance for evaluating PAHs in soil at LANL PRSs. ER personnel are cautioned that evaluations completed according to this guidance may require revision at a later date. It is highly recommended that a risk assessor and statistician from the Assessments Council be consulted before implementing any of the procedures described in this document.

References:

- Bradley, L.J.N., B.H. Magee, and S.L. Allen. 1994. Background Levels of Polycyclic Aromatic Hydrocarbons (PAH) and Selected Metals in New England Urban Soils. Journal of Soil Contamination, 3(4):349-361.
- Los Alamos National Laboratory. 1995. Statistical Comparisons to Background, Part I, Los Alamos National Laboratory, Environmental Restoration Project, Assessments Council, March 28, 1995 (LA-UR-95-1217).

Appendix A **DRAFT**

The Environmental Restoration Project at the Los Alamos National Laboratory has worked with its regulators at Region VI of the Environmental Protection Agency to determine an appropriate methodology for comparing analytical concentration data to local background concentrations. The methods that have been agreed upon are detailed in the policy paper, *Statistical Comparisons to Background, Part I, Los Alamos National Laboratory, Environmental Restoration Project, Assessments Council*, (LA-UR-95-1217). One of the methods set forth in that paper involves comparison with an upper tolerance limit (UTL) of the background distribution. The UTL used is the 95th percent upper confidence limit on the 99th percentile of the estimated distribution.

The data provided by Bradley, et al., included numbers of samples, numbers of detections, minimum observed values, maximum observed values, arithmetic means, geometric means, and 95th percent upper confidence limits on the arithmetic means. The confidence limits were calculated assuming normality of the underlying distribution. If this assumption were true, the arithmetic means would be expected to fall in the middle of the range described by the minimum and maximum. In actuality, the means fall far closer to the minimum than the maximum in every case, showing that the data were skewed to the right and not normally distributed. Given the apparent shape of the distribution, lognormal models were selected to describe the data sets for the purpose of calculating UTLs.

A lognormal distribution can be described by two parameters that must be estimated in order to simulate the data. The geometric mean provides a slightly positive-biased estimator for one of the parameters. However, no estimator of the second parameter was available from the Bradley data. In order to estimate the second parameter, simulations were performed. First, sample variances, s^2 , were back-calculated from the 95th percent upper confidence limits on the arithmetic means (given in Bradley, et al.) by subtracting the mean, dividing by 1.96 (as used by Bradley et al.), multiplying by n (62), and squaring the results. One thousand data sets of size sixty two were then simulated from lognormal distributions with the first estimated parameter held constant. The second parameter was varied until the average sample variance, s^2 , of the one thousand data sets matched the sample variance, s^2 , observed by Bradley, et al. The parameters of the lognormal distributions were then recorded.

For each compound, a new set of simulations was run to calculate the UTL based on estimated lognormal distributions. The simulations included four hundred data sets of size one thousand. From each set of observations, the 99th percentile was identified. The averages of the 99th percentiles from the four hundred data sets were calculated and recorded as estimates of the 99th

DRAFT

percentiles for each distribution. The four hundred 99th percentiles were combined in a data set and ranked. The non-parametric estimates of the 95th percent upper confidence limit on the 99th percentiles (UTLs) were estimated as 97.5th percentiles from these data sets.

The UTL was reported for every compound that had less than 25% non-detections in the original data. When the number of non-detections was greater than 25%, the maximum observed value was reported.

Statistical Comparisons to Background

Part I

Los Alamos National Laboratory

**Environmental Restoration Project
Assessments Council**

March 28, 1995

INTRODUCTION

The main purpose of this policy is to provide guidance to the Los Alamos National Laboratory (the Laboratory) Environmental Restoration (ER) personnel on recommended statistical comparisons between Potential Release Site (PRS) data and naturally occurring concentrations of metals (including radionuclides) in soils. Any background comparison approach consists of two components. The first is to assemble a defensible set of background data. This document describes comparisons to the Laboratory wide set of background data collected by Longmire (Longmire et al. 1994, 1142). The second is the selection of the statistical method(s) used to compare site data to background data. Two statistical methods are presented. In the first method, site concentration data are compared with a statistic representing the upper percentile of background concentrations, the upper tolerance limit (UTL). The UTL is used as a screening value, or extreme (hot) value, to determine if a significant release has occurred at the site. The second is a group of methods designed to detect a distributional shift between site data and background data. Although guidelines for the application of these methods are presented in this document, each data analysis report should briefly describe the statistical analysis method chosen and justify its application to the data in question. In PRS data, particular attention should be paid to background comparisons of beryllium and arsenic, because background concentrations of these elements exceed risk-based screening action values. Each data analysis report should also justify the use of Laboratory-wide background concentration data, or present the rationale for using site-specific background concentration data.

Comparisons of PRS data with background concentrations are needed as part of the Resource Conservation and Recovery Act (RCRA) corrective action process. The guidance on statistical comparisons between PRS data and naturally occurring concentrations of metals (including radionuclides) in soils provided by this document may also be applicable to background comparisons of certain anthropogenic compounds (e.g., radionuclides distributed from nuclear fallout or organic chemicals associated with urban activities). It should be emphasized that the approach presented in this document should be applied to naturally-occurring metals only.

REGULATORY LITERATURE REVIEW

The Environmental Protection Agency (EPA) guidance documents supporting the Comprehensive Environmental Response Compensation and Liability Act (CERCLA) and RCRA

programs provide specific information on how to design background studies and how to statistically compare site data with background data.

The CERCLA document, *Guidance on Data Useability in Risk Assessment (Part A)* (EPA 1992, 1166), recommends collecting background data prior to collecting site data. If the comparison of background data with site-derived data for a given constituent does not show a difference statistically, that constituent is eliminated from further analyses. The CERCLA guidance also suggests basing the number of background samples collected from a site on the "minimum detectable difference" procedure (EPA 1989, 0303). Data analysts unfamiliar with this approach should contact the statistical specialists designated by the Assessments Council Chair.

Background comparisons for groundwater monitoring data are addressed in the RCRA document, *The RFI Guidance* (EPA 1989, 0088). Methods for comparing data derived from upgradient wells with data from downgradient wells is presented in the RCRA groundwater statistical analysis document (EPA 1989, 1141), referenced in the EPA RCRA facility investigation (RFI) guidance (EPA 1989, 0088). These statistical methods are codified in 40 CFR Part 264, *Statistical Methods for Evaluating Ground-Water Monitoring from Hazardous Waste Facilities: Final Rule* Federal Register Tues. Oct. 11, 1988.

Statistical methods used for background comparisons of groundwater can be applied to background comparisons for data from other media as stated in the preface of EPA (1989, 1141):

"This scenario can be applied to other non-RCRA situations involving the same spatial relationships and the same null hypothesis. The explicit null hypothesis for testing contrasts between means, or where appropriate between medians, is that the means between groups (here monitoring wells) are equal (i.e., no release has been detected), or that the group means are below a prespecified action level (e.g., the ground-water protection standard). Statistical methods that can be used to evaluate these conditions are described in Section 5.2 (Analysis of Variance), 5.3 (Tolerance Intervals), and 5.4 (Prediction Intervals)." Bold added for emphasis.

The RCRA groundwater monitoring guidance suggests that the specific approach proposed by the owner/operator should be submitted to EPA for approval, especially where methods other than those presented in the guidance are used. Statistical methods presented below are consistent with those found in the analysis of variance and tolerance interval sections of the RCRA groundwater monitoring document (EPA 1989, 1141).

BACKGROUND COMPARISON APPROACHES AT OTHER DOE FACILITIES

Based on information presented at the 1994 Technical Information Exchange (TIE) Workshop, most DOE facilities have funded a facility-wide background analysis of all potentially impacted media (soil, sediment, surface water, groundwater). Most have information on the soil horizon or other data that support a site-specific comparison with background. No single statistical test for comparing site data with background data is used throughout the DOE facilities studied.

Background comparison approaches at specific facilities include:

- 1) Hanford Site (Richland, WA): a background conceptual model has been developed. This model includes metals (including certain radionuclides) and considers transport between soil, sediment, surface water and groundwater.
- 2) Oak Ridge Reservation (Oak Ridge, TN): a statistically significant difference between site data and background data is required before including the constituent in a risk assessment.
- 3) Sandia National Laboratories (Albuquerque, NM): upper tolerance limits (UTL) of metals and radionuclides are calculated based on historical data collected within and outside of PRSs. UTL calculations are made if there is an adequate number of "detects" reported by the analytical laboratory. The UTL is calculated using a method dependent on the distributional properties of the analyte.

LABORATORY BACKGROUND DATA

The strategy at the Laboratory has been to collect samples that are representative of Laboratory-wide background metal concentrations in soil and tuff (see Longmire et al. 1994, 1142) for comparison with PRS data. (Readers interested in more detail on the statistical distribution of naturally-occurring metals are referred to the Appendix of this paper.) Longmire's data, representing Laboratory-wide variation in soil and volcanic tuff, are used as the default background data for making comparisons in the initial RFI screening assessment process. The Laboratory-wide background data were collected from sites representing the range of soil conditions observed at the Laboratory. At present, Longmire has 47 soil samples (A, B or C soil horizon) analyzed by EPA SW 846 methods, and 50 soil (A, B or C soil horizon) and 38 tuff samples analyzed by Instrumental Neutron Activation Analysis (INAA). The INAA data represent

total elemental concentrations and are useful for background comparisons only at selected PRSs with data using a comparable method of analysis. Additional soil, sediment and volcanic tuff background analyses will be added to the Laboratory-wide background data base during 1995. The new data will be made available to ER Project personnel and other interested parties as soon as validated data are received.

Variation in certain elements makes comparisons with Laboratory-wide background less valid than comparisons with site-specific background concentrations. For example, due to natural variability in element concentrations of Bandelier Tuff, background soils from Technical Areas (TAs) at the east and west ends of the Laboratory are likely to be enriched, or depleted, in certain elements relative to Laboratory-wide background values. Bandelier Tuff was derived from a zoned magma chamber in which some elements were concentrated at the top and others were concentrated at the bottom. During the eruption resulting in deposition of the Bandelier Tuff, magmas at the top of the chamber were erupted first. Consequently, those elements that were concentrated in the upper magmas are in higher abundance at TAs to the east of the Laboratory because the Bandelier Tuff subunit exposed in that region represents the earlier-erupted magmas. Thus, soils derived from tuff located to the east of the Laboratory are likely to have higher abundance of certain elements than Laboratory-wide background soils. For example, at TA-33 uranium in background soils is at higher levels than Laboratory-wide background because it is more abundant in the stratigraphically lowest units of the Tschirege Member of the Bandelier Tuff. Thus, it is important to develop site-specific background for some sites.

All users of background data must follow some simple guidelines to ensure that site data is being compared with Laboratory background data in a scientifically valid manner. A primary requirement is that PRS data and background data be collected from soil or tuff having the same physical properties. If PRS data are collected from sediments, background data should also come from sediments. For example, if site data were collected exclusively from a soil horizon naturally enriched in a metal, these data would appear to be greater than Laboratory background data. A second requirement is that PRS samples be assayed by the same analytical methods as used for background samples.

Each data analysis report should demonstrate that the above guidelines were considered before using Laboratory-wide background distributions. Failure to meet these guidelines may be an indication that site-specific background be collected or that a subset of Laboratory-wide background data be compared with PRS data. Lack of data for a particular analyte could be justification for the collection of site-specific background data. Before recommending the

collection of site-specific background for an analyte, data analysts should confirm whether the analyte in question could be present by looking at historical information. On the other hand, lack of defensible historical information might justify the collection of site-specific background data.

Should site-specific background data seem warranted, data analysts are encouraged to seek technical advice from the background specialists designated by the Laboratory ER Project Assessments and Earth Sciences Council Chairs. These specialists will be able to provide detailed guidance on the location and number of samples required for site-specific background comparisons.

PROPOSED STATISTICAL METHODS

Because background comparisons are used to make decisions throughout the RCRA process from site screening to corrective measures implementation, data analysts must have statistical methods that can be applied over a broad range of decisions. This guidance defines two statistical methods for background comparisons. Both methods meet the requirements for RCRA decision-making. In the first method, the "hot measurement" test, site concentration data are compared with a statistic representing the upper percentile of background concentrations. In the second method, the "distributional shift test," the mean of site data is compared with the mean of background data to determine if the former is statistically greater than the latter. Used together, or separately, these tests help demonstrate if a release at a PRS occurred, and help define what risk consequence the release may have. Figure 1 illustrates the differences between site data and background data detected by the two methods.

The decision to be supported by the background comparison determines which test is more appropriate. In an initial RFI screening, when a single high value should trigger further analysis, the hot measurement test is typically more appropriate. When extensive data are collected to support a risk assessment and a change in the average concentration should lead to further action at the site, the distributional shift test may be more appropriate. A data analysis report should clearly indicate the rationale for selecting a statistical method that differs from those presented in this guidance document.

It is emphasized that the level of effort spent in evaluating potential differences between PRS data and background data should be related to the site-specific information available. For example, if historical information indicates that beryllium was released at a firing site, the potential

differences between the beryllium concentration data at the firing site and Laboratory or site-specific background data should be carefully evaluated and presented in the data analysis report.

COMMON DATA ANALYSIS PROCEDURES

Because certain methods depend on the statistical distributions of site and background data, analysts are encouraged to prepare graphical data displays to facilitate the communication of the results of data comparisons. Box plots (see Figure 1), in which background data and site data can be compared side-by-side, are most useful. Analysts should also consider using histograms and probability plots. These graphs provide tangible evidence of the similarity or differences between site data and background data.

HOT MEASUREMENT TEST

The hot measurement test defines a threshold value that represents high background concentrations. No matter what parameters are chosen to define the threshold, there exists a probability that a background measurement will exceed the hot measurement threshold. The frequency of false positive results is minimized by using a threshold statistically related to higher background concentrations. The confidence limit on a percentile of the distribution, termed the tolerance limit, is such a value, and is one of the background comparison methods recommended by EPA (1989, 1141). The Laboratory has selected the 99th percentile for calculating the upper tolerance limit, based on the general guidance in the RCRA groundwater document. If the underlying distributional model is correct, the upper tolerance limit based on the 99th percentile is rarely exceeded. EPA recommends calculating an upper 95% confidence limit for the target percentile (EPA 1989, 1141). The Upper Tolerance Limit (UTL) for the 99th percentile at 95% confidence can be calculated using Equation (1).

$$UTL_{0.99,0.95} = \text{mean} + \text{standard deviation} * k_{0.99,0.95} \quad (1)$$

The "k-factor" depends on the number of background samples, and complete tables of k-factors are published in EPA (1989, 1141) and Gilbert (1987, 0312). Table 1 presents k-factors selected to represent the range of values used to compute UTLs for Laboratory background soil samples. To apply Equation (1), the background data must be normally distributed or transformed to normality (e.g., by using log-transformation). If data deviates sufficiently from normality, nonparametric methods for calculating tolerance limits should be considered (e.g., as described in

Gilbert [1987, 0312]). Alternatively, when appropriate, the data analyst may trim outliers from the distribution and calculate the UTL based on the trimmed mean and standard deviation.

The Laboratory soil background data were used to calculate the UTL values presented in Table 2. As discussed in section 4.0, use of the default is emphasized, or Laboratory-wide background data should be justified for the specific background comparison being performed. Table 2 UTL values should not be used without considering the guidelines discussed in section 4.0. A minimalist approach to data preparation was used for the initial UTL calculations. Because some soil concentration data were sufficiently skewed, a log-transformation was applied to improve the fit of these data to a normal distribution. (Readers interested in the details of these data distributions are referred to the Appendix of this policy paper.) The UTLs calculated for the lognormal distribution were backtransformed into original units to simplify comparisons with site data. No values were trimmed from the distributions used to calculate the mean and standard deviation. If four or fewer non-detects were reported for an analyte, values below the laboratory detection limit were replaced by one-half of the detection limit (EPA 1989, 1141). The UTL was not calculated for any analyte having more than four non-detect values (>10% non-detects).

The observed maximum concentration in the background data is an alternative to using the UTL as the hot measurement threshold. However, when few background samples are available, using the maximum will result in an underestimation of the upper background percentile. In general, the sample maximum (for "n" samples) is an estimate of the $\left[\frac{n-0.5}{n} \right]$ th percentile. Thus, if 10 samples are collected, the sample maximum is an estimate of the 95% percentile. As more Laboratory-wide background data become available, the maximum value will increase, but the UTL will typically not change. Because the maximum is extremely sensitive to background sample size, it is not recommended for use as a hot measurement threshold. Rarely detected analytes, which include: antimony, cadmium, mercury, selenium and thallium, are an exception to this general recommendation. For this limited subset of rarely detected analytes, the maximum detected background concentration can be used as the hot measurement threshold.

Exceeding the UTL does not prove that a release occurred at a PRS. Assuming the PRS is at background and the statistical model is correct, there is a 1% probability that the 99th percentile will be exceeded by each sample collected from the PRS. Furthermore, a typical metals suite requires comparison of 23 analytes to background. If the concentrations of the 23 metal analytes vary independently, the 1% probability that each PRS sample exceeds the 99th percentile increases to a 21% probability that at least one of the 23-99th percentiles will be exceeded in a

single sample. Additionally, given that the probability values for these multiple comparisons have not been adjusted, the overall level of confidence for 23 analytes will be substantially less than 95%. Consequently, the results of a hot measurement test must be carefully evaluated. The possibility of exceeding a UTL due to an unusual, but naturally occurring, soil matrix is a further consideration.

The results of the UTL comparison should also be evaluated relative to potential human health or ecological screening levels. Some analytes, arsenic and beryllium in particular, represent a special case. Because background levels at the Laboratory exceed risk-based screening levels, no screening action levels (SALs) for beryllium or arsenic have been calculated for the Laboratory ER Program. Seven metals (antimony, barium, chromium, lead, manganese, thallium, and vanadium) and one radionuclide (thorium-232) have background UTLs that appear close to their SAL values. The UTLs of this group represent a significant fraction of the SAL (8 to 50%). If, in a comparison similar to a multiple constituent test, each of these metals had a concentration equal to the SAL, the total of the metal UTLs divided by the corresponding SAL (in effect normalizing the UTL to SAL ratio) would equal 89%. Thus, the concentrations of most naturally-occurring metals are significantly lower than their respective SALs.

Both the multiple constituent evaluation and the UTL-to-SAL comparison will help determine what level of effort should be expended to evaluate deviations from background. For most naturally-occurring metals, when only a single statistical comparison to background is performed, the UTL will be adequate because probability levels are not compromised. Under this circumstance, the UTL is the simplest comparison and is functionally most similar to comparisons of site data to target risk levels or SALs.

DISTRIBUTIONAL SHIFT TEST

The distributional shift test is used to determine if site data are systematically greater than background data. Several types of distributional shift tests are available. The Student t-test is a parametric, statistical, two-sample test that determines whether the mean concentration of site data is statistically greater than the mean concentration of background data. The Wilcoxon rank-sum test is the nonparametric equivalent to the t-test (Gilbert 1987, 0312; Gilbert and Simpson 1992, 0974). The Wilcoxon test pools site and background data into one aggregate set and determines if the average rank of site data is greater than that of the background data. The Wilcoxon test is recommended when site data consists of few samples or when non-detects are frequent. Another useful distributional shift test is the Quantile test (Gilbert and Simpson 1992,

0974). This test, which compares the upper quantile (e.g., 25%) of background data with that of PRS data, is more capable of detecting a difference when only a small number of PRS concentrations are elevated. The Quantile test is the most useful distributional shift test for PRSs at which samples from a release represent a small fraction of the overall data collected at the PRS because it does not artificially reduce statistical significance. For example, to detect contamination from historical spills at unknown locations, an RFI work plan may call for samples to be collected from a grid. Most sample results show no contamination, but those in or near spill locations show elevated concentrations.

Use of the distributional shift test is dependent on the number of samples available for comparison. In general, at least 10 sample concentrations for comparison with background data are needed to provide adequate confidence for detecting a shift. Frequently, in RFI Phase I, inadequate numbers of samples are collected to warrant a distributional shift comparison.

To infer a significant result in a distributional shift test, a 95% confidence level is recommended. Given that multiple comparisons will be performed with the distributional shift test, the same statistical interpretation issues cited above for the hot measurement test are also present. In addition, the human health and ecological consequences of a PRS concentration data above background must be considered along with differences in metal concentrations between soil horizons. In particular, multiple comparison tests with SALs must be performed and ecological SAL comparisons must be made.

REFERENCES

EPA (US Environmental Protection Agency) 1989. "Ecological Assessment of Hazardous Waste Sites: A Field and Laboratory Reference," EPA/600 3-89/013, Environmental Research Laboratory, Corvallis, OR. (EPA 1989, 0303)

EPA (US Environmental Protection Agency) 1989. "Interim Final RCRA Facility Investigation (RFI) Guidance, Volume I of IV, Development of an RFI Work Plan and General Considerations for RCRA Facility Investigations," EPA/530-SW-89-031, OSWER General Directive 9502.00-6D, Office of Solid Waste, Washington, DC. (EPA 1989, 0088)

EPA (US Environmental Protection Agency) 1989. "Statistical Analysis of Ground-Water Monitoring Data at RCRA Facilities. Interim Final Guidance," Office of Solid Waste, Waste Management Division, US Environmental Protection Agency, Washington DC. April 1989. (EPA 1989, 1141)

EPA (US Environmental Protection Agency) 1992. "Guidance for Data Useability in Risk Assessment (Part A)," Office of Emergency Remedial Response, US Environmental Protection Agency, Washington DC. April 1992. (EPA 1992, 1166)

Gilbert, R.O. 1987. Statistical Methods for Environmental Pollution Monitoring. Von Nostrand Reinhold Company Inc. New York. 320 pp. (Gilbert 1987, 0312)

Gilbert, R. O., and J. C. Simpson 1992. "Statistical Methods for Evaluating the Attainment of Cleanup Standards, Volume 3: Reference-Based Standards for Soils and Solid Media," Pacific Northwest Laboratory, Richland, Washington 99352. (Gilbert and Simpson 1992, 0974)

Longmire, P., S. Reneau, P. Watt, J. Gardner, and C. Duffy. 1994. Geomorphology, Pedogenesis, and Geochemistry of Background Bandelier Tuff and Selected Soil Profiles, Los Alamos, New Mexico, 1994. (Longmire et al. 1994, 1142)

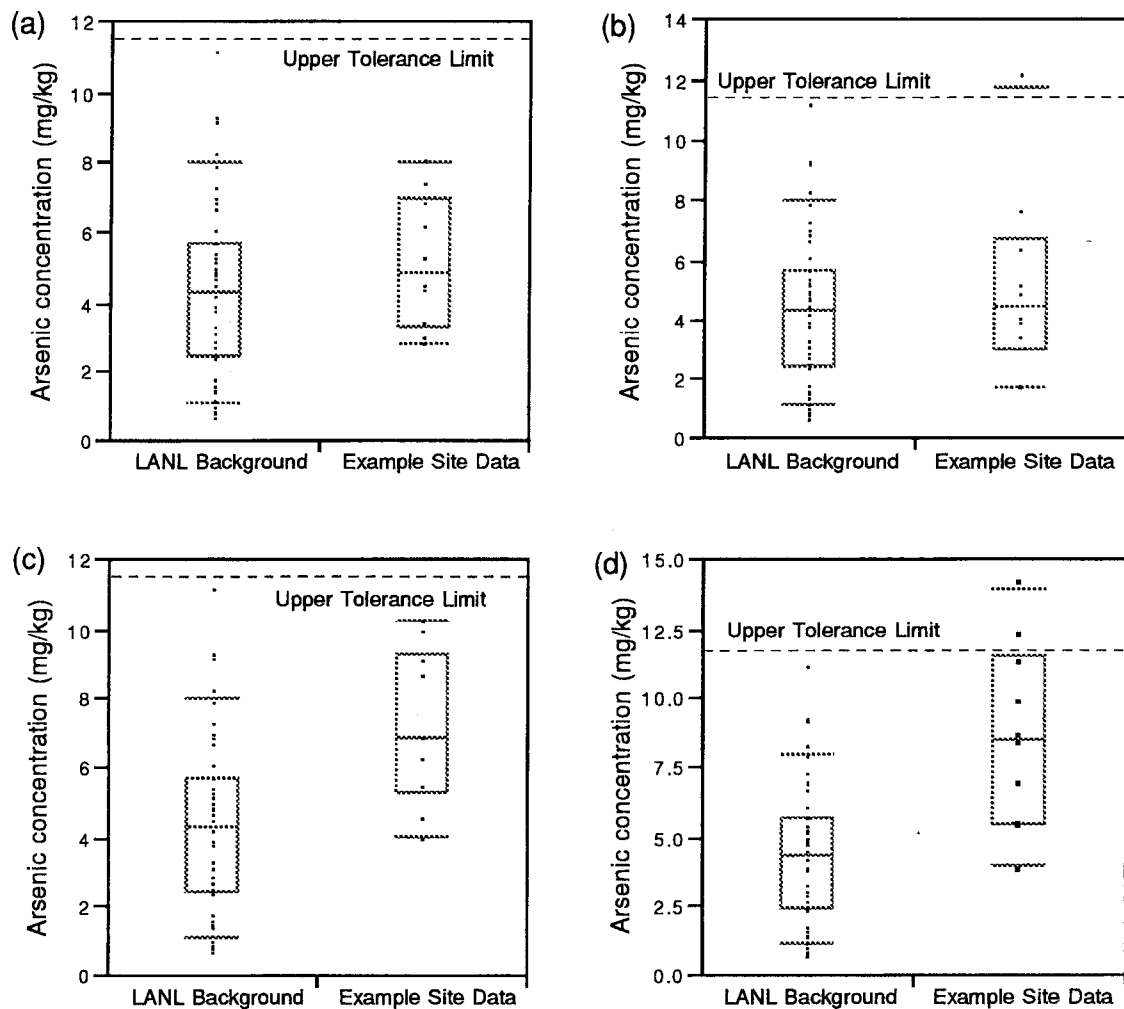


Figure 1. Box plot comparisons of example site data to Laboratory background data.
 (a) Site data are within the range of background: no distributional shift or hot measurements (i.e., values greater than the UTL).
 (b) Site data fail hot measurement test: one of 11 site arsenic concentration values exceeds the UTL of 11.6 mg/kg.
 (c) Site data show a distributional shift: the Wilcoxon rank sum test shows that site data tends to be greater than the background data.
 (d) Site data show a distributional shift and fail the hot measurement test: two of 10 site arsenic concentrations exceed the UTL of 11.6 mg/kg and the site data tend to be greater than the background data.

Statistical Comparisons to Background, Part I

Table 1 - Selected "k-factors" used to calculate UTLs (* reprinted from Gilbert 1987, 0312).

Number of background samples	$k_{0.99,0.95}$
45	2.897 *
46	2.8902
47	2.8834
48	2.8766
49	2.8698
50	2.863 *

Table 2 - List of UTLs for LABORATORY soil background data.

Analyte	SAL (mg/kg)	Mean (mg/kg)	Standard deviation (mg/kg)	UTL _{99%,0.95} (mg/kg)	N	N > DL *
Aluminum (LT)		19000	13800	123000	47	47
Antimony	32	2.45	0.36	2.5 (MAX)	46	2
Arsenic		4.4	2.5	11.6	46	46
Barium (LT)	5,600	161	129	1140	47	47
Beryllium		1.15	0.75	3.31	47	47
Cadmium	80	0.39	0.54	2.7 (MAX)	47	5
Calcium (LT)		5790	12500	54400	47	47
Chromium (Total) †		11.7	7.8	34.2	47	47
Cobalt (LT)		15.2	7.6	51.1	47	47
Copper	3,000	5.3	3.6	15.7	47	45
Iron		14500	7320	35600	47	47
Lead	400	15.0	8.3	39.0	47	44
Magnesium (LT)		2920	2150	16100	47	47
Manganese	11,000	343	238	1030	47	47
Mercury	24	0.05	0.01	0.1 (MAX)	48	4
Molybdenum	400	NA	NA	NA	NA	NA
Nickel	1,600	9.7	5.9	26.7	47	45
Potassium		2420	1304	6180	47	47
Selenium	400	0.43	0.41	1.7 (MAX)	46	23
Silver	400	NA	NA	NA	NA	NA
Sodium (LT)		577	453	3320	47	47
Strontium	48,000	NA	NA	NA	NA	NA
Thallium	6.4	0.27	0.24	0.9 (MAX)	45	21
Vanadium	560	25	14	66	47	47
Zinc	24,000	41	21	101	47	47
⁴⁰ Potassium (1)		21.6	5.07	36.1	50	50
²³² Thorium (1)	5	1.71	0.34	2.68	50	50
²³⁴ Uranium (1)	86	1.21	0.29	2.03	50	50
²³⁵ Uranium (1)	18	0.052	0.012	0.088	50	50
²³⁸ Uranium (1)	59	1.14	0.27	1.90	50	50

* Concentration values < detection limit (DL) were replaced by 1/2 of the DL

† - SAL for Chromium-III is 80,000 mg/kg and for Chromium-VI is 400 mg/kg

LT - UTL is based on log transformed data

NA - data not available for Laboratory background

MAX - Maximum value is reported, rather than the UTL

(1) - Data are converted from elemental concentrations reported in the Laboratory background report. Units are in pCi/g.

Background Levels of Polycyclic Aromatic Hydrocarbons (PAH) and Selected Metals in New England Urban Soils

L. J. N. Bradley,^{1*} B. H. Magee,² and S. L. Allen¹

¹ENSR Consulting and Engineering, 35 Nagog Park, Acton, MA 01720

²Ogden Environmental and Energy Services, 239 Littleton Road, Suite 7C, Westford, MA 01886

* To whom all correspondence should be addressed.

ABSTRACT: Polycyclic aromatic hydrocarbons (PAH) are byproducts of combustion and are ubiquitous in the urban environment. They are also present in industrial chemical wastes, such as coal tar, petroleum refinery sludges, waste oils and fuels, and wood-treating residues. Thus, PAHs are chemicals of concern at many waste sites. Risk assessment methods will yield risk-based cleanup levels for PAHs that range from 0.1 to 0.7 mg/kg. Given their universal presence in the urban environment, it is important to compare risk-based cleanup levels with typical urban background levels before utilizing unrealistically low cleanup targets. However, little data exist on PAH levels in urban, nonindustrial soils. In this study, 60 samples of surficial soils from urban locations in three New England cities were analyzed for PAH compounds. In addition, all samples were analyzed for total petroleum hydrocarbons (TPH) and seven metals. The upper 95% confidence interval on the mean was 3 mg/kg for benzo(a)pyrene toxic equivalents, 12 mg/kg for total potentially carcinogenic PAH, and 25 mg/kg for total PAH. The upper 95% confidence interval was 373 mg/kg for TPH, which exceeds the target level of 100 mg/kg used by many state regulatory agencies. Metal concentrations were similar to published background levels for all metals except lead. The upper 95% confidence interval for lead was 737 mg/kg in Boston, 463 mg/kg in Providence, and 378 mg/kg in Springfield.

KEY WORDS: background, PAH, metals, urban, anthropogenic, soil.

1. INTRODUCTION

Polycyclic aromatic hydrocarbons (PAHs) are byproducts of combustion and are naturally occurring chemicals in the environment. Forest fires and volcanoes are major natural sources of PAHs, but there are anthropogenic sources as well due to burning of fossil fuels, including automobile and industrial emissions. PAHs are chemicals of concern in many waste site investigations that are undertaken pursuant to the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA), the Resource Conservation and Recovery Act (RCRA), and state

hazardous waste programs. Risk assessments performed according to federal guidance for former manufactured gas plant sites, wood treating facilities, petroleum refineries, and other sites generally conclude that PAHs pose unreasonable risks to human health and that remedial actions must be taken to reduce risks to acceptable levels. The majority of the risk posed by PAHs is generally due to benzo(a)pyrene and the other PAHs that have been shown to cause cancer in laboratory animals after repeated dosings. The U.S. EPA (1993a) currently identifies seven PAHs as "probable human (B2) carcinogens": benzo(a)pyrene, benzo(a)anthracene, benzo(b)fluoranthene, benzo(k)fluoranthene, chrysene, dibenz(a,h)anthracene, and indeno(1,2,3-c,d)pyrene.

Because of the very health-protective assumptions used in regulatory risk assessments, very low risk-based clean-up levels for PAHs are derived for such sites. In Michigan, residential soil cleanup levels of 0.33 mg/kg for each carcinogenic PAH have been set (MDNR, 1993). In New Jersey, proposed residential soil cleanup levels are 0.66 mg/kg for benzo(a)pyrene (New Jersey Register, 1992). The use of standard CERCLA risk assessment guidance (U.S. EPA, 1993b) results in the derivation of a risk-based cleanup level for benzo(a)pyrene of 0.1 mg/kg.

All of these risk-based soil cleanup levels are below the urban, nonindustrial background soil concentrations presently reported in the literature. However, the availability of such data is very limited. Blumer (1961) reports that benzo(a)pyrene concentrations in Cape Cod, MA, soils range from 0.04 to 1.3 mg/kg. Menzie et al. (1992) report that urban background soil levels of total carcinogenic PAH range from 0.06 to 5.8 mg/kg. Butler et al. (1984) report that total PAH levels in soils alongside roadways in England range from 4 to 20 mg/kg, and potentially carcinogenic PAH range from 0.8 to 11.5 mg/kg. Blumer et al. (1977) report that total PAH levels in soils in a Swiss town range from 6 to 300 mg/kg.

It is very difficult to compare the data from these studies to the results of site risk assessments due to the limited dataset and the nonuniformity of the PAH compounds evaluated. Clearly, more data are required from nonindustrial urban locations to define the urban background level for PAH and to critically evaluate the role of risk assessment in setting remedial goals for PAH in soils. Accordingly, we have collected 60 samples of surficial soils from urban locations in three New England cities and analyzed them for all 17 PAH compounds present on the EPA's Target Compound List, which is used in the Superfund program. In addition, all samples were analyzed for total petroleum hydrocarbons (TPH) and for seven metals: arsenic, barium, cadmium, chromium, lead, mercury, and selenium.

II. METHODS

A. Sample Collection

Samples of surficial soils from urban locations in three New England cities were collected: Boston, MA; Providence, RI; and Springfield, MA. Twenty independent

samples and duplicates of two samples were collected in each city. The samples were collected on July 21, 22, and 23, 1992, respectively. The samples were taken at a depth of 0 to 6 in. in areas considered to be not directly affected by industrial sites. Generally, the locations were along roads and sidewalks, and in parks and open lots. Each location was characterized in writing, including a soil description, and photographically documented. The samples were collected following standard environmental sampling protocols (U.S. EPA, 1986).

B. Sample Analysis

Chemical analysis of the samples was performed by AnalytiKEM, Inc. (Cherry Hill, NJ). The samples were analyzed by GC-MS for the 17 PAH compounds present on the EPA's Target Compound List using the methods required by EPA Method 8270 for the analysis of semivolatile compounds. In addition, the samples were analyzed for the eight RCRA metals, total petroleum hydrocarbons (TPH; EPA Method 418.1), and total solids. The complete analyte list is given in Table 1.

C. Data Validation

Validation of the data received from AnalytiKEM was performed according to U.S. EPA (1991) guidelines. The data were reviewed for completeness, holding times, GC-MS tuning and system performance, initial and continuing calibrations, laboratory method blank analysis, surrogate recoveries, matrix spike and matrix spike duplicate analysis, field duplication precision, and compound quantitation and detection limits.

D. Data Analysis

The analytical data were summarized in accordance with U.S. EPA (1989) risk assessment guidance. If a compound was detected at least once in surface soil, one half the sample quantitation limit (SQL) was used as a proxy concentration for all samples reported as "below detection limit" in the estimation of exposure point concentrations. However, if a compound was not detected in any sample, that compound was omitted from further consideration. In addition, when a proxy concentration (i.e., one half the detection limit) was greater than the highest actual detected value for a compound in any sample, that concentration was considered to be an aberration and was omitted from the database. This is consistent with U.S. EPA (1989) guidance, which recognizes that high sample quantitation limits can lead to unrealistic concentration estimates.

TABLE 1
Chemical Analyses of Urban Soils

Semivolatile Organics, EPA Target Compound List

Naphthalene
Acenaphthylene
Acenaphthene
Fluorene
Phenanthrene
Anthracene
Fluoranthene
Pyrene
Benzo(a)anthracene
Chrysene
Benzo(b)fluoranthene
Benzo(k)fluoranthene
Benzo(a)pyrene
Indeno(1,2,3-cd)pyrene
Dibenzo(a,h)anthracene
Benzo(g,h,i)perylene
2-Methylnaphthalene

Metals

Arsenic, total
Barium, total
Cadmium, total
Chromium, total
Lead, total
Mercury, total
Selenium, total
Silver, total

Other

Total petroleum hydrocarbons
Solids

A slightly different method of analysis was used to evaluate PAH. Because PAH are generally found in groups, it was conservatively assumed that if one PAH was detected in a sample, other compounds in that class might also be present in that sample. Therefore, if one PAH was detected in a sample, all undetected PAH were assigned a proxy concentration equal to one half the SQL. If a sample had no detected PAH, no PAH were assumed to be present in the sample, and a concentration of zero was used for all nondetects.

Summary statistics (minimum, maximum, arithmetic mean, upper 95% confidence limit on the arithmetic mean, and frequency of detection) were generated for each compound for each city and for all three cities combined.

The data for PAH were summarized in several different ways. Of the 17 PAH analyzed in each sample, seven are considered to be probable human carcinogens (Group B2) by the U.S. EPA (1993a). The U.S. EPA has derived a cancer slope factor, which is a measure of the carcinogenic potency of a compound, only for benzo(a)pyrene (B(a)P) (U.S. EPA, 1993a). Review of the literature indicates that not all PAH are equally potent with respect to tumor induction. Several researchers have proposed toxic equivalency schemes that relate the tumorigenic potency of each PAH to that of B(a)P (ICF-Clement Associates, 1988; Woo, 1989). B(a)P toxic equivalency factors (B(a)P-TEFs) can be used to adjust either the B(a)P dose-response value to provide a compound-specific dose-response value, or the concentration of each PAH in a sample to be expressed in terms of B(a)P toxic equivalents (B(a)P-TE). The latter method was used here. B(a)P-TE were calculated using the B(a)P toxic equivalency factors recommended for use by the U.S. EPA (1993c), as shown in Table 2. For each sample, PAH concentrations were reported for each of the 17 PAH on the analyte list, for total PAH (tPAH), for total carcinogenic PAH (cPAH), and for B(a)P-TE, and these values were used to generate the summary statistics for each group of samples.

III. RESULTS

Analysis of the laboratory results for the PAH indicates that quality control criteria were acceptable. The data were analyzed to determine if any statistically significant differences existed between the datasets for the three cities. A Hartley test for homogeneity of variances (Mendenhall, 1979) and a one-factor analysis of variance to test for equality of the means (Mendenhall, 1979) indicated no statistically significant differences. The results indicate that the PAH data can be pooled and treated as one dataset for further statistical analyses.

TABLE 2
Benzo(a)Pyrene Toxic
Equivalent Factors (BAP-TEF)

Compound	EPA TEF
Benzo(a)pyrene	1.0
Benz(a)anthracene	0.1
Benzo(b)fluoranthene	0.1
Benzo(k)fluoranthene	0.1
Chrysene	0.001
Dibenzo(a,h)anthracene	1.0
Indeno(1,2,3-c,d)pyrene	0.1

The results of the PAH analyses are presented in Table 3 for all cities combined. A summary of the PAH results by city and for all cities combined is presented in Table 4, which reports for each: tPAH, total cPAH, and total B(a)P-TE. The arithmetic mean and the upper 95% confidence limit concentration are reported for each. Table 4 provides a summary of the data by city, and the results are graphically presented in Figure 1.

Table 5 presents a summary of the metals, TPH, and solids data by city. A Hartley test for homogeneity of variances and a one-factor analysis of variance to test for equality of the means indicated that the metals and TPH data from the three cities cannot be combined. This is due to the fact that the concentrations in each city are not normally distributed and did not have equal variances. The concentrations of the metals are compared to the arithmetic mean concentrations in the eastern U.S. (ATSDR, 1992) in Table 5. Most notably, lead concentrations are much higher than background concentrations. This is most likely due to the effects of automobile exhaust.

In order to determine if sample location significantly affected PAH concentration results, individual samples were classified based on the sample location's

TABLE 3
Summary Statistics for PAH — All Areas Combined

Compound	Minimum detect (mg/kg)	Maximum detect (mg/kg)	Arithmetic mean	Upper 95% interval (mg/kg)	Frequency of detection ^a	
2-Methylnaphthalene	0.017	0.64	0.151	0.173	19	62
Acenaphthene	0.024	0.34	0.201	0.306	30	62
Acenaphthylene	0.018	1.10	0.173	0.208	24	62
Anthracene	0.029	5.70	0.351	0.535	54	62
Benzo(a)anthracene	0.048	15.00	1.319	1.858	58	62
Benzo(a)pyrene	0.040	13.00	1.323	1.816	57	62
Benzo(b)fluoranthene	0.049	12.00	1.435	1.973	55	62
Benzo(g,h,i)perylene	0.200	5.90	0.891	1.195	36	62
Benzo(k)fluoranthene	0.043	25.00	1.681	2.522	59	62
Chrysene	0.038	21.00	1.841	2.693	60	62
Dibenzo(a,h)anthracene	0.020	2.90	0.388	0.521	32	62
Fluoranthene	0.110	39.00	3.047	4.444	60	62
Fluorene	0.022	3.30	0.214	0.317	35	62
Indeno(1,2,3-c,d)pyrene	0.093	6.00	0.987	1.293	43	62
Naphthalene	0.018	0.66	0.125	0.149	35	62
Phenanthrene	0.071	36.00	1.838	2.982	61	62
Pyrene	0.082	11.00	2.398	2.945	61	62
Total BAP-TE	0.257	21.31	2.437	3.324	62	62
Total carcinogenic PAH	0.680	77.70	8.973	12.423	62	62
Total PAH	2.292	166.65	18.361	24.819	62	62

^a Frequency of detection = number detected: number samples.

TABLE 4
Background PAH Concentrations in Urban Surface Soils^a

Compound	Boston (n = 20)		Providence (n = 20)		Springfield (n = 20)		All cities (n = 60)	
	Arithmetic mean (ppm)	Upper 95% CI (ppm)	Arithmetic mean (ppm)	Upper 95% CI (ppm)	Arithmetic mean (ppm)	Upper 95% CI (ppm)	Arithmetic mean (ppm)	Upper 95% CI (ppm)
Total B(a)P-TB	2.4	4.6	2.1	2.9	2.8	4.5	2.4	3.3
Total cPAH	8.4	16.0	7.8	11.0	10.6	18.3	9.0	12.4
Total PAH	18.7	35.9	16.8	23.5	19.1	29.9	18.4	24.8
TPH	474.9	652.6	267.4	338.2	184.4	233.3	306.2	372.8

^a 0 to 6 in.

355

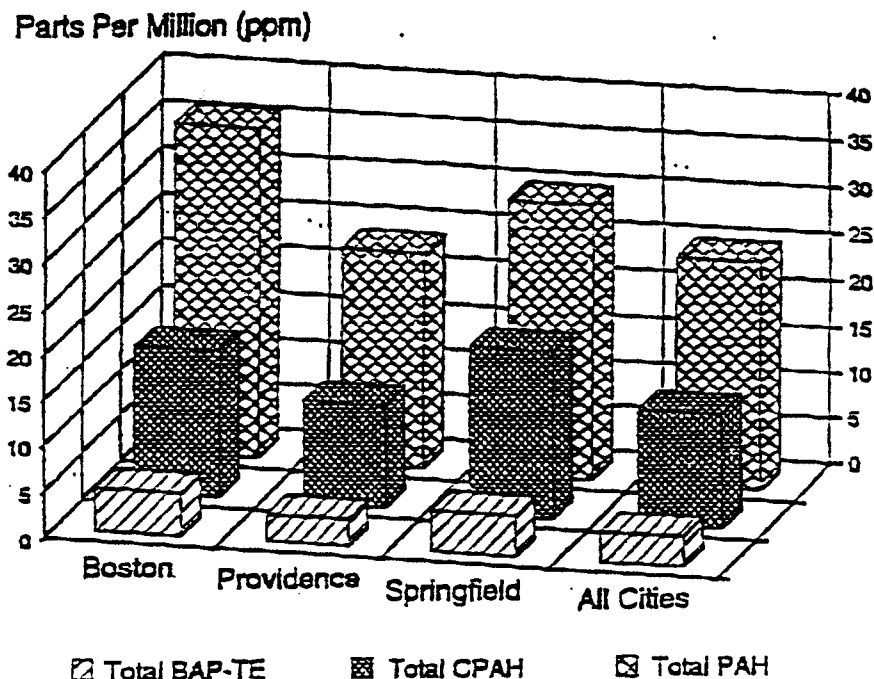


FIGURE 1. Background concentrations of PAH in urban soils. Data presented are the upper 95% confidence interval on the arithmetic mean. Data are presented numerically in Table 4.

proximity to asphalt pavement, based on both written and photographic documentation of sample location. Generally, samples collected within 4 to 6 ft of a road were considered to be near pavement. Of the 60 separate locations, 42 were considered to be near pavement and 18 were not. When tested for equality of variance and means as above, the two populations were determined to be significantly different. The mean total PAH concentration near pavement was 22 ppm compared to 8 ppm not near pavement. These results are shown in Table 6.

Similar analyses were performed to see if TPH or total organic carbon concentrations could be used as surrogates for PAH concentrations. The results showed that there is no correlation between PAH and TPH concentrations, nor between PAH and total organic carbon concentrations (data not shown).

The highest total PAH concentration detected was 166 mg/kg, taken from a street corner in Boston. The next highest PAH concentration was 109 mg/kg, taken at the base of a telephone pole. Four of the 60 samples were taken at the bases of telephone poles, with widely varying results. The total PAH concentrations in the other three locations were 62, 4, and 45 mg/kg.

TABLE 5
Summary Statistics for Metals, TPH, and Soils by City

Compound	Boston (n = 20)		Providence (n = 20)		Springfield (n = 20)		Arithmetic mean in U.S. soils ^a (mg/kg)
	Arithmetic mean (mg/kg)	Upper 95% Interval (mg/kg)	Arithmetic mean (mg/kg)	Upper 95% Interval (mg/kg)	Arithmetic mean (mg/kg)	Upper 95% Interval (mg/kg)	
Arsenic, total	4.20	5.59	3.53	4.27	5.63	9.23	7.4
Barium, total	53.95	66.25	45.29	59.43	45.17	51.03	420
Cadmium, total	1.55	2.79	ND	ND	ND	ND	0.25 ^b
Chromium, total	23.00	27.69	12.08	14.35	12.62	14.45	52
Lead, total	398.70	737.44	305.76	462.98	261.69	377.76	17
Mercury, total	0.29	0.39	0.19	0.24	0.20	0.25	0.12
Selenium, total	0.51	0.57	0.39	0.48	0.53	0.55	0.45
Total petroleum hydrocarbons	474.90	652.62	267.43	338.19	184.38	233.27	—
Total solids	90%	93%	93%	95%	90%	92%	—

^a ATSDR. 1992. *Public Health Assessment Guidance Manual*. PB92-147164. U.S. Department of Health and Human Services.

^b ATSDR. 1991. *Toxicological Profile for Cadmium*. PB92-147164. Draft. U.S. Department of Health and Human Services.

TABLE 6
Comparison of Background PAH Concentrations in Urban Soils: The Effects of Proximity to Pavement

Compound	Results of statistical analysis									
	Near pavement		Not near pavement		Test for homogeneity of variances			Test of equality of means		
	Arithmetic mean (ppm)	Standard deviation	Arithmetic mean (ppm)	Standard deviation	Sample F-statistic	Associated degrees of freedom	Statistically significant at 0.05 level of significance	Sample Student's t	Associated degrees of freedom	Statistically significant at 0.05 level of significance
Total B(a)P-TH	2.9	4.1	1.1	0.92	21.3	41, 17	Yes	2.69	50	Yes
Total PAH	21.9	30.7	8.3	7.2	18.4	41, 17	Yes	2.69	50	Yes

IV. CONCLUSION

In this study, 20 surface soil samples were collected from each of three New England cities and analyzed for PAH, TPH, and metals. The results of the statistical analyses described in the previous section show that, with respect to PAH, the three datasets are not significantly different and can be considered as one dataset representative of urban environments. The samples were taken in typical urban areas but not near known industrial sites. Therefore, these data are considered to be representative of the generalized effects of urban activities.

It is clear from the results presented here that common regulatory target cleanup levels for cPAH and B(a)P-TE (0.1 to 0.66 mg/kg) are much below the background concentrations of these compounds in urban surface soils (upper 95% confidence interval of 3.3 and 12.4 mg/kg for total B(a)P-TE and total cPAH, respectively). Figure 2 graphically compares the "bright line" target cleanup level for B(a)P of 0.1 mg/kg with the total B(a)P-TE (upper 95% confidence interval on the arithmetic mean) measured in urban environments.

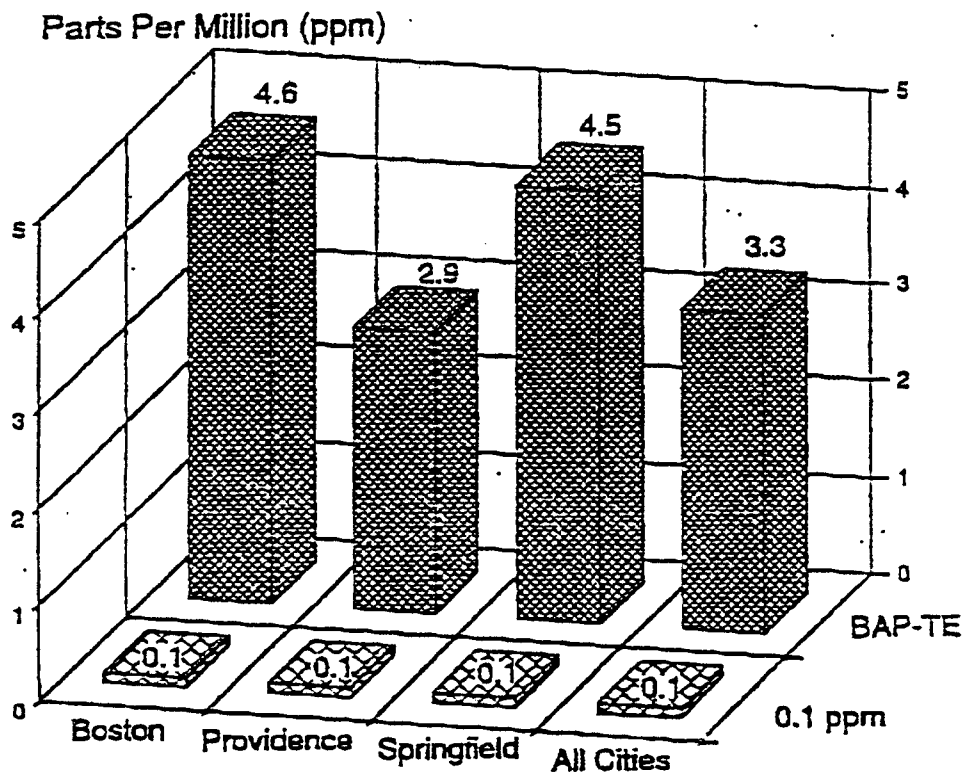


FIGURE 2. Comparison of B(a)P-TE with U.S. EPA Region III⁴ risk-based concentration for B(a)P. B(a)P data presented are the upper 95% confidence interval on the arithmetic mean.

Upper 95% confidence intervals are compared because this is the statistic preferred by EPA and many states for risk assessment. Moreover, the State of Massachusetts defines its background concentrations of metals based on the upper 95% confidence limit on the arithmetic mean concentration (Massachusetts Department of Environmental Protection, 1992). For all cities combined, the background level of B(a)P-TE of 3.3 mg/kg is approximately ten times greater than the target cleanup level of 0.33 mg/kg and approximately 30 times higher than the target cleanup level of 0.1 mg/kg. For those regulatory situations in which the use of B(a)P-TEFs in determining site risk is not allowed, the background level of cPAH is approximately 40 to 100 times greater than these target cleanup levels.

An analysis of the data comparing samples taken near pavement with those determined to be not near pavement indicated that those samples designated near pavement had significantly higher, approximately threefold higher, PAH concentrations for both total PAH and total B(a)P-TE. This is most likely due to the presence of diesel and automobile exhaust particles, perhaps influenced by the presence of asphalt and runoff of vehicular oil from the roads.

Total petroleum hydrocarbons (TPH) were also found at consistently high levels in each city. The commonly applied regulatory cleanup level for TPH is 100 mg/kg. This cleanup level is not risk based and is three times lower than the background concentration of TPH found in this study (arithmetic mean of 306 mg/kg and upper 95% confidence interval on the mean of 373 mg/kg).

It is incumbent upon the regulatory agencies to recognize that substantial background levels of PAH and TPH exist in our urban environments and to acknowledge this information in the development of realistic target cleanup levels. The use of these background data in setting more realistic target cleanup levels may result in better allocation of remedial and regulatory dollars in site investigations.

REFERENCES

-
- Agency for Toxic Substances and Disease Registry, (ATSDR). 1991. *Toxicological Profile for Cadmium*. Draft. U.S. Department of Health and Human Services.
- Agency for Toxic Substances and Disease Registry (ATSDR). 1992. *Public Health Assessment Guidance Manual*. PB92-147164. U.S. Department of Health and Human Services.
- Blumer, M. 1961. Benzpyrenes in soil. *Science*. 134:474.
- Blumer, M., Blumer, W., and Reich, T. 1977. *Polycyclic Aromatic Hydrocarbons in Soils of a Mountain Valley: Correlation with Highway Traffic and Cancer Incidence*. Work at Woods Hole supported by the Office of Naval Research (N00-14-66 Contract CO-241) and the National Science Foundation (Grant DES 74-22781).
- Butler, J. D. et al. 1984. Some observations on the polycyclic aromatic hydrocarbon (PAH) content of surface soils in urban areas. *Sci. Total Environ.* 33:75.
- ICF-Clement Associates. 1988. *Comparative Potency Approach for Estimating the Cancer Risk Associated with Exposure to Mixtures of Polycyclic Aromatic Hydrocarbons*. Prepared for U.S. EPA Office of Health and Environmental Assessment. (Interim Final Report) Contract No. 68-02-4403.

- Massachusetts Department of Environmental Protection. 1992. *Risk Assessment Short Form*. Boston, MA, Office of Research and Standards and the Bureau of Waste Site Cleanup, Massachusetts Department of Environmental Protection.
- Mendenhall, W. 1979. *Introduction to Probability and Statistics*. Fifth ed. Belmont, CA, Duxbury Press.
- Menzie, C. A., Potocki, B. B., and Santodonato, J. 1992. Exposure to carcinogenic PAHs in the environment. *Environ. Sci. Technol.* 26:1278.
- Michigan Department of Natural Resources. 1993. *MZRA Operational Memorandum #8, Revision 2*. July 16.
- New Jersey Register. 1992. *Site Remediation Program. Cleanup Standards for Contaminated Sites. Proposed New Rules: NJAC 7:26D. C24:373*. Monday, February 3.
- U.S. EPA. 1986. *SW446*. Vol. 2, 3rd ed. Washington, D.C., OSWER, U.S. Environmental Protection Agency.
- U.S. EPA. 1989. *Risk Assessment Guidance for Superfund. Volume I: Human Health Evaluation Manual (Part A)*. EPA 540/1-89/002. Washington, D.C., Office of Emergency and Remedial Response, U.S. Environmental Protection Agency.
- U.S. EPA. 1991. *National Functional Guidelines for Organic Data Review*. Draft. June. U.S. Environmental Protection Agency.
- U.S. EPA. 1993a. *Integrated Risk Information System (IRIS)*. Cincinnati, OH, Environmental Criteria and Assessment Office, U.S. Environmental Protection Agency.
- U.S. EPA. 1993b. *Risk-Based Concentration Table, Third Quarter*. U.S. Environmental Protection Agency, Region III.
- U.S. EPA. 1993c. *Provisional Guidance for Quantitative Risk Assessment of Polycyclic Aromatic Hydrocarbons*. EPA/600/R-93/089. U.S. Environmental Protection Agency.
- Woo, Y. T., Health and Environmental Review Division, Office of Toxic Substances, U.S. Environmental Protection Agency, Washington, D.C. 20460. June 14, 1989. *PAHs Screening for Carcinogenicity*. Memorandum to P. Tong, Hazardous Site Evaluation Division, Office of Emergency and Remedial Response, U.S. Environmental Protection Agency, Washington, D.C. 20460.