

# **Kentucky Guidance for Ambient Background Assessment**

**January 8, 2004**



**Natural Resources and  
Environmental Protection Cabinet**

## **Introduction**

This guidance document is intended to assist in comparing site data and background data for sites undergoing environmental assessment. These procedures provide a simplified statistical procedure for determining if the site data is part of the background population. It also provides generic statewide background values for inorganic chemicals that may be used in lieu of collecting site-specific background samples. The statistical procedures may be used for site-specific data or the generic statewide values in Tables 1 and 2. This guidance does not preclude other appropriate statistical comparisons from being made, but rather a simplified screening method that does not require a deep knowledge of statistics. If the site data set fails the statistical procedures in this guidance, it may be appropriate to perform a more complete statistical comparison.

Background, as defined in 401 KAR 42:005 (definitions codified to support the Underground Storage Tank regulations), means the concentration of substances consistently present in the environment at, or regionally proximate to, a release but outside the influence of the release. There are two types of background:

- a) Natural background is the amount of naturally occurring substances in the environment, exclusive of that from anthropogenic sources.
- b) Ambient background means the concentrations of naturally occurring inorganic substances and ubiquitous anthropogenic inorganic substances in the environment that are representative of the region surrounding the site and not attributable to an identifiable release.

Since sites undergoing environmental assessment are often found in industrialized and potentially contaminated areas, the determination of site-specific background concentrations is difficult. Generic ambient background values applicable to all sites in Kentucky would be useful for comparison to site data for the purpose of identifying those constituents requiring remedial action (i.e., removal or exposure control). These generic ambient background values would provide an alternative to attempting to identify site-specific background soils in areas that are likely contaminated.

## Methodology

To provide an alternative to site-specific background sampling, the NREPC used background sample values provided by regulated facilities, as well as background sample values collected by cabinet employees. These samples were collected from areas generally considered to be outside of the influence of site activities, but were potentially impacted by regional or urban activity. Therefore, these samples represent “ambient,” as opposed to “natural,” background. From 400 to over 800 samples for each constituent were used in the analysis. For each constituent, a 95% Upper Confidence Limit (UCL) of the arithmetic mean, 60<sup>th</sup> percentile, and 95<sup>th</sup> percentile were calculated. The 95% UCL is the value below which the true mean of the data set falls, with 95% confidence. The 60<sup>th</sup> and 95<sup>th</sup> percentiles indicate that 60 percent and 95 percent of the data falls below those values.

The following methodology was employed to calculate ambient background:

1. Values reported as “non-detected” were retained in the database at half the reporting limit (USEPA, 1998).
2. As the data sets came from areas having varied uses (e.g., industrial, commercial, residential, agricultural, woodlands, etc.), the probability that some of the samples were taken in contaminated areas is significant. Data sets were tested for outliers by the Grubb’s test, and individual samples that had a calculated Z-score above 3.8 were generally removed from the background data set. The Grubb’s test formula is as follows:

$$Z = \frac{|population\ mean - value\ of\ individual\ sample|}{standard\ deviation}$$

3. The descriptive statistics of mean and standard deviation were calculated by standard parametric methods assuming normality and are listed in Table 1. Parametric methods were used to allow for comparisons between these generic ambient background values and the results of other published studies of background.

a. Standard deviation was calculated by the “nonbiased” method employing the formula:

$$S.D. = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n-1}}$$

b. Mean was calculated as the sum of all individual scores divided by the total number of observations.

4. The data sets were analyzed with Lillefor’s test for normality. Since the data sets are not normally or lognormally distributed, the parameters that are to be used in determining if site samples are consistent with background (i.e. 95% UCL of mean, 60<sup>th</sup> percentile and 95<sup>th</sup> percentile) were calculated by nonparametric methods and are listed in Table 2.
5. The 95% UCL of the arithmetic mean for each constituent was calculated on the trimmed data set using ProUCL. ProUCL is a statistical package developed by Lockheed Martin under contract with the U.S. EPA.
6. The 60<sup>th</sup> percentile value is used as the midpoint for each constituent. It was calculated as follows:
  - a. The constituent values were ranked in increasing order of magnitude.
  - b. The quantity  $60(n)/100$  was used to identify the measurement with the resulting rank.
7. The 95<sup>th</sup> percentile value is used as the upper bound value for each constituent and was calculated as follows:
  - a. The constituent values were ranked in increasing order of magnitude.
  - b. The quantity  $95(n)/100$  was used to identify the measurement with the resulting rank.

The thallium data were characterized by a large number of non-detects (633 non-detects verses 54 detects). Due to the large number of non-detects, non-detects were not entered as  $\frac{1}{2}$  the non-detect concentration. Each non-detect sample was assumed to have a concentration equal to the recorded non-detect concentration. Considering the number of non-detects and the likelihood that

the recorded values skew thallium concentrations upward, only the 95<sup>th</sup> percentile of the total data is cited in Table 2.

### **Procedure for Comparison to Background**

The site data should be segregated by surface and subsurface data. The surface and subsurface site data may be compared to the statewide numbers in Table 2, or to site-specific background samples. The following three criteria may be used to demonstrate that the site data is background:

1. The mean site concentration for inorganic constituents must be below the 95% UCL of the mean concentrations of background for inorganic constituents.
2. At least half of the data points should be less than the 60<sup>th</sup> percentile.
3. No data points should be above the upper bound value (95<sup>th</sup> percentile).

These procedures provide a tool for comparing site data with either generic statewide or site-specific background using the statistical characteristics of the two populations. Other statistical comparisons may be used, if appropriate.

### **Determining Site-specific Background**

Site-specific ambient background levels may be determined at the site. The site-specific ambient background data set shall consist of an appropriate number of samples for the statistical method employed. The number of samples necessary to characterize site-specific background will vary based on the variability of the data. Twenty data points may be used as a minimum number of samples per horizon (surface and subsurface) as a default number, unless other statistical methods can be used to develop a different number. A site-specific determination of the number of required samples may be calculated based on the statistical characteristics of the background population.

Upgradient groundwater samples are to be obtained from the same hydrogeological unit as the groundwater contamination at the site. The background monitoring wells shall be located hydrogeologically upgradient from the release(s) of concern, unless it can be demonstrated to the cabinet that the upgradient location is undefinable or infeasible.

Background soil samples should be collected from native soil in areas of similar soil type as found at the site. Background concentrations should be determined separately for surface and subsurface areas that are consistent with the on-site investigation.

The following areas are inappropriate to sample when determining soil background unless otherwise necessary to reach a corrective action decision or identify potential sources of contamination:

1. Fill areas;
2. Areas in which management, treatment, handling, storage or disposal activities of any of the following are known or suspected to have occurred: hazardous substances or petroleum, solid or hazardous wastes, or waste waters;
3. Areas within three feet of a roadway;
4. Parking lots and areas surrounding parking lots or other paved areas;
5. Railroad tracks or railway areas or other areas affected by their runoff;
6. Areas of concentrated air pollutant depositions or areas affected by their runoff;
7. Storm drains or ditches presently or historically receiving industrial or urban runoff;  
or
8. Areas within three feet of any current structure, or the former location of any structure, which is likely to have been painted with lead-based paint.

## Literature Cited

United States Environmental Protection Agency (USEPA), 1995. Determination of Background Concentrations of Inorganics in Soils and Sediments at Hazardous Waste Sites. Office of Research and Development. Office of Solid Waste and Emergency Response. EPA/540/S-96/500. December, 1995.

United States Environmental Protection Agency (USEPA), 1998. Statistical Tests for Background Comparison at Hazardous Waste Sites. Supplemental Guidance to RAGS: Region 4 Bulletins – Addition #1. Interim Draft. USEPA Region 4, Waste Management Division. Atlanta, Georgia. November, 1998.

Table 1. Summary Statistics for Ambient Inorganic Chemicals

<b>Element</b>	<b>Number of Samples</b>	<b>Range (mg/kg)</b>	<b>Mean (mg/kg)</b>	<b>Standard Deviation (mg/kg)</b>
Aluminum	679	1290 - 38,100	10969	5462.9
Arsenic	539	0.059 - 55.5	8.9	7
Barium	756	6.14 - 1160	111.3	92.4
Beryllium	696	0.061 - 3.57	0.8	0.5
Cadmium	701	0.004 - 9.46	0.68	1.4
Chromium	771	2.83 - 168	20.5	13.9
Cobalt	649	0.29 - 67.6	11.9	8.1
Copper	729	0.49 - 636	18.9	39.7
Iron	697	222 - 86,900	22456	13269.7
Lead	808	0.03 - 284	30	31.3
Manganese	685	8.43 - 5100	1017	854.9
Mercury	459	0.007 - 0.721	0.06	0.1
Nickel	716	0.39 - 83.7	20.9	13.1
Selenium	714	0.001 - 3.93	0.94	0.7
Silver	697	0.006 - 5.2	0.42	0.6
Thallium	633	0.13 - 28		
Vanadium	679	4.82 - 92.1	26.9	11.8
Zinc	721	6 - 470	55	46.3



Table 2. Generic Statewide Ambient Background for Kentucky

<b>Element</b>	<b>Mean (mg/kg)</b>	<b>95% UCL of Mean (mg/kg)</b>	<b>60<sup>th</sup> Percentile (mg/kg)</b>	<b>95<sup>th</sup> Percentile (mg/kg)</b>
Aluminum	10969	11314	10800	21000
Arsenic	8.9	9.4	8.3	21.2
Barium	111.3	116.9	100	241
Beryllium	0.8	0.83	0.75	1.8
Cadmium	0.68	0.78	0.27	3.9
Chromium	20.5	21.3	19.3	40
Cobalt	11.9	12.4	13.1	25.1
Copper	18.9	21.3	13.8	41.7
Iron	22456	23284	22000	47600
Lead	30	33	20.9	84.6
Manganese	1017	1071	948	2620
Mercury	0.06	0.07	0.059	0.14
Nickel	20.9	21.7	20.2	46.8
Selenium	0.94	0.99	1.38	2.1
Silver	0.42	0.45	0.257	1.2
Thallium				7.95
Vanadium	26.9	27.7	27.3	48.6
Zinc	55	57	48.6	115