Use of Statistics for Determining Soil/ Groundwater Cleanup Levels under the Risk Reduction Rules

Effective Date: April 30, 1998

This memo outlines the acceptable use of statistical approaches on site data, specifically the upper confidence limit (UCL) for comparison to a cleanup level. The use of a UCL for demonstrating compliance with a cleanup level is prescribed in 30 Texas Administrative Code, Chapter 335.553(d) of the Risk Reduction Rules (RRRs). The RRRs do not state any limitation for the use of the UCL other than a minimum data set of 10 samples or provide any guidance for the appropriate use of statistics on various media and exposure scenarios. The following sections discuss the appropriateness of using a UCL for various media and exposure pathways within the context of the RRR in the Voluntary Cleanup Program (VCP). The memo also briefly discusses the development of background concentrations using statistics and the comparison of site data to background.

SOIL PROTECTIVE OF INGESTION AND INHALATION

As described in the U.S. Environmental Protection Agency (EPA) supplemental guidance to their Risk Assessment Guidance for Superfund (Calculating the Concentration Term, Vol. 1, No. 1) exposure to site contaminants over a long (chronic) period of time is best represented by an arithmetic average concentration. Because an individual is assumed to move randomly across an exposure area over time, the spatially averaged soil concentration can be used to estimate the true average contaminant concentration contacted over time. Therefore, it is appropriate to compare the 95% UCL of a set of soil data to a medium specific concentration (MSC) that is protective of soil ingestion/inhalation. Please be aware that site-specific exposure scenarios may exist that are not appropriately modeled using this statistical technique. In these situations, other methodologies should be incorporated to best approximate the conditions at the site.

GROUNDWATER

A UCL statistical analysis of groundwater for comparison to a cleanup level can be allowed as long as appropriate criteria are followed. According to the EPA's April 1989 document entitled *"Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities, Section 6.2.1"*, "A confidence interval for the mean concentration is constructed from the sample data for each compliance well individually. These confidence intervals are compared with the compliance limit." In other words, a 95% UCL can be determined for each well at the site, assuming that at least 10 discrete samples have been analyzed from each well, which then can be compared to the Risk Reduction Standard (RRS) 2 cleanup level. Calculating a UCL from groundwater analytical data obtained from various wells located throughout the contamination plume is not an appropriate statistical analysis. The fact that exposure to groundwater occurs at immovable locations (i.e., water wells) makes it inappropriate to average groundwater contamination across a site.

SOIL PROTECTIVE OF CROSS-MEDIA CONTAMINATION (GROUNDWATER)

The soil level that is protective of groundwater (Soil - GWP) has been designed to ensure soil contamination below the Soil - GWP will not impact groundwater in excess of the RRS 2 MSC. Soil contamination exceeding the Soil - GWP or an adjusted Soil - GWP is deemed to be a threat to groundwater. Therefore, direct comparison of discrete soil contamination levels to the Soil - GWP value is the appropriate method for determining groundwater protection. A statistical analysis of soil contamination averaged throughout the extent of the impacted soil cannot demonstrate compliance with a Soil - GWP since any specific soil contamination level exceeding a Soil - GWP is a threat to groundwater.

In summary, the use of a site-averaged UCL for comparison to an established cleanup level is only appropriate for the soil ingestion/inhalation pathway. Direct comparison to a cleanup level is appropriate for Soil-GWP. The calculation of a UCL for comparison to a groundwater cleanup level is only appropriate for data collected from individual well; averaging analytical data from two or more wells is not acceptable.

STATISTICS AND BACKGROUND

An accepted statistical method for determining a background value from a set of data is the 95% upper tolerance limit (UTL). The UTL represents a value that 95% of the population will fall below with 95% confidence. Usually, the UTL will tend to be higher than the highest value in the background data set that was used to calculate the UTL. An adequate number of samples should be collected (at minimum, 8 samples per lithology or zone of concern) for the UTL to represent site background conditions. Any single data point from the site that exceeds the background UTL indicates that contamination is present. Developing a statistically-derived, averaged value from site data for comparison to the background UTL is not acceptable. In other words, only single data points are to compared to the background UTL.

The UTL can be calculated as follows:

  - Calculate the mean (x) and the standard deviation (S) from the data set.

  - Construct the one-sided upper tolerance limit as:
        UTL = x + kS;
        k = one-sided tolerance factor (see Table I).

  - Compare each analytical value from the area of concern to the UTL. Any data point exceeding the UTL is statistically significant evidence of contamination.

Table I

TOLERANCE FACTORS (k) FOR ONE-SIDED NORMAL TOLERANCE
INTERVALS WITH PROBABILITY LEVEL (CONFIDENCE FACTOR) Y= 0.95 AND
COVERAGE P = 95%)

n	k
3	7.655
4	5.145
5	4.202
6	3.707
7	3.399
8	3.188
9	3.031
10	2.911
11	2.815
12	2.736
13	2.670
14	2.614
15	2.566
16	2.523
17	2.486
18	2.543
19	2.423
20	2.396
21	2.371
22	2.350
23	2.329
24	2.309
25	2.292
30	2.220
35	2.166
40	2.126
45	2.092
50	2.065

n = number of samples in background data set.

SOURCE: Lieberman, Gerald F., 1958. "Tables for One-sided Statistical Tolerance Limits." *Industrial Quality Control*. Vol. XIV, No. 10.