

APPROPRIATE STATISTICAL METHODS  
FOR THE COMPARISON OF REMEDIAL INVESTIGATION DATA  
TO BACKGROUND DATA  
AT THE ROCKY FLATS PLANT SITE

The draft final Phase III RFI/RI Report for Operable Unit 1 presents two different statistical approaches for the evaluation of chemicals and radionuclides detected in the remedial investigation against background concentrations as determined in the background geochemical characterization study. In the public health evaluation the information used to select the correct statistical method included the probability density function of the sample population, the sample population variance and standard deviation, and the sample population mean. Based on this information a parametric or nonparametric statistical analysis was selected and conducted in a manner consistent with generally accepted statistical procedures. This comparison has been generally referred to in recent comment resolution meetings as inferential statistics or an ANOVA approach. In contrast a second statistical approach was applied in the remedial investigation and the environmental evaluation sections of the report for the same background evaluation. For these sections a simple comparison was made between the maximum site specific concentration and the calculated upper tolerance limit (UTL) of the background sample population from the background geochemical study.

The ANOVA and UTL approaches have different purposes and applications. Consequently they can result in conflicting conclusions even when applied to the same data sets. The selection of the most appropriate statistical methodology for the RI report should be based on the purpose of the evaluation. In this case the purpose is to determine whether site specific chemical and radionuclide levels are significantly greater than background levels of those constituents.

It is EPA's position that the ANOVA approach is the appropriate statistical methodology to use for this purpose. We believe this approach must be consistently applied to the evaluation in the remedial investigation, the public health evaluation, and the environmental evaluation. Our rationale for taking this position can be summarized as follows:

- a. As a comparison technique between sample sets of remedial investigation data and background data sets, the ANOVA is a more powerful method. Inherently comparisons of mean differences are less sensitive than comparisons of individual values against some background data set (DOE's UTL approach). This is chiefly because mean value comparisons like ANOVA are less sensitive to modest departures from an assumed distribution.
- b. Results and conclusions based on the ANOVA approach can be easily verified.

c Weaknesses of the UTL approach include the fact that it is sensitive to the number of comparisons made. Obviously a few outlier data comparisons with a background data set is different from the hundreds of sample comparisons contemplated by DOE. There are a variety of fairly involved statistical techniques to handle multiple comparisons of individual data. EPA RCRA guidance provides for factoring in multiple comparisons (such as the Bonferroni approach) with ANOVA. DOE is apparently attempting to address this problem with the UTL by using the 5% rule (sample data sets are at background unless more than 5% of the samples exceed the UTL). However this rule is not sensitive to the absolute values of the exceedances. EPA's concern is that a single value 100 times the UTL is more important than one just 10% above yet the distinction will apparently not be made if the 5% rule is met for a sample data set. As presently understood the 5% rule does not appear to account for the randomness inherent in samples drawn from some assumed background population. Literature references on prior uses of this method and mathematical validation for the approach have not been provided by DOE.

d The UTL can not be used to determine whether there is a statistical difference between populations. It only compares individual maximum concentrations with the calculated UTL. No other statistical information such as the distribution of the data, the mean concentration or standard deviation is used to carry out the comparison. Thus it is possible that the site related mean contaminant concentration is significantly different from the background population mean but fewer than 5% of the samples exceed that UTL. The result may be an erroneous conclusion that the remedial investigation sample population is no different from the background sample population. In this case the standard deviation for the site related contaminants would be less than the standard deviation for background. It is also possible that comparisons of data with different statistical distributions will be made using the UTL approach.