CHAPTER 15 PROGRAM NOTE #15:
STATISTICAL SOFTWARE PACKAGES APPLICABLE FOR USE AT MSW LANDFILLS, WITHOUT SITE-SPECIFIC REVIEW AND ACCEPTANCE November 10, 1994

On or before October 9, 1993, each of nine Regional Water Quality Control Boards (**RWQCB**s) passed their own version of a "**Super Order**" which amended the waste discharge requirements (**WDR**s) of all municipal solid waste landfills (**MSW landfills**) in their Region to bring those WDRs into compliance both with Chapter 15 and with the "**federal MSW regulations**" [40 CFR Part 258]. Such rapid statewide action was required by State Water Resource Control Board (**SWRCB**) Resolution No. 93B62 [entitled Policy For Regulations Of Discharges Of Municipal Solid Waste (**Policy**)], and helped facilitate the United State's Environmental Protection Agency's determination that California's regulatory program for MSW landfills is equivalent to the federal MSW regulations. This "approved state" status provides the RWQCBs with greater flexibility in regulating MSW landfills.

The Super Order contained a statistical/non-statistical data analysis section. Most MSW landfills will begin using this in October of 1994. This Program Note introduces computer programs that various companies have developed to implement the data analysis portion of the Super Order. The programs are listed in the order in which they were determined to be applicable for such use. This Program Note will be revised whenever either: (**a**) a new data analysis package is deemed applicable (only one program is approved, so far, but others are currently under review); or (**b**) one of the previously addressed programs is revised to offer new capabilities that can be used without additional site-specific review.

The use of such pre-evaluated statistical software packages by the RWQCBs and the regulated community should result in a savings of both time and resources.

PACKAGE #1: GSAS, Version 1.3

APPLICABILITY

Staff of the SWRCB and the Board's statistical consultant, Dr. Neil Willits of the Statistical Laboratory, University of California at Davis, have reviewed Version 1.3 of the Groundwater Statistical Analysis System (GSASTM), a computer program by **Intelligent Decision Technologies, Ltd.** ["**IDT**", 3308 Fourth St., Boulder, Colorado 80304 // (303) 449B2457] and find that it is applicable for use at all MSW landfills in California that are subject to the Super Order; however, such use is applicable only under the conditions listed below. Three methods are pre-accepted and three methods need site-specific acceptance.

From a computational standpoint, **GSAS** meets the basic requirements both of the federal MSW regulations and of Article 5 of Chapter 15. It should be recognized that there are some analysis methods that would also be acceptable but which are not implemented in Version 1.3 of GSAS-two such methods are mentioned in the following discussion. Most of the following discussion describes the ways in which **this** version of **GSAS** can be

used and mentions various additional features the software will have available in future (as yet unreviewed) versions.

DISCUSSION

Version 1.3 of **GSAS** implements two pre-accepted statistical methods and one pre-accepted non-statistical method, in addition to three other methods that need site-specific justification. All statistical methods are subject to the considerations discussed in the section labeled "CHOICE OF A STATISTICAL METHOD".

METHODS IN GSAS V-1.3

Methods Pre-accepted: The following analyses within **GSAS** are applicable without the need for substantiation by the discharger: They can be used in detection monitoring without the need for a review of the appropriateness of the software on a facility-by-facility basis:

Parametric ANOVA;

Nonparametric ANOVA-In addition to its standard usage under the Super Order, this test can be used in lieu of a Test of Proportions (when the Super Order requires the Test of Proportions), provided the test is based on the sample size required for the Test of Proportions; and

California Non-statistical VOC Analysis.

Methods Needing Site-specific Acceptance: One of the three following methods within **GSAS** may be more appropriate in its implemented form than the statistical analyses incorporated within the Super Order; however, the use of these methods requires facility-specific justification. Therefore, although **GSAS** properly applies these methods, their use is subject to site-specific review and acceptance by the RWQCB:

Prediction Limits;

Control Charts; and

Tolerance Intervals.

CHOICE OF A STATISTICAL METHOD:

The statistical analysis of groundwater data can be conceived as consisting of two parts. First, a method (or family of methods) must be chosen for the analysis, and then the analysis associated with that method needs to be carried out.The choice of a method is frequently done on a one-time basis, though it is good practice to review the validity of the assumptions that are required for the analysis on a continuing basis. One of the strengths of **GSAS** is that once a family of methods (such as ANOVA or Prediction Intervals) has been chosen, it will choose an appropriate analysis variant from within that family. For example it will choose between a parametric analysis, an analysis of transformed data, and a nonparametric analysis, based on an assessment of whether the data seem to be normally distributed with constant residual variance.Prior to this point, however, the choice of the family of methods must be justified. This justification would ordinarily occur prior to the point when statistical analysis of the collected data is initiated, or else whenever a change in the method of analysis is proposed. The justification must address seasonality, trends in background, appropriateness of control

sites (i.e., the source of "background" data), and independence of samples, as discussed below.

Seasonality: An analysis of groundwater data can be based upon data that are all collected on the same sampling date, over a small range of dates, or it can involve data that are collected over a more extensive range of dates. (Under the blanket Super Order, the samples for an ANOVA should be taken within a 30 day period.) If the proposed analysis involves data collected over an extended time range (i.e., involving a period longer than 30 days), then the data should be checked statistically to determine if seasonal patterns are present. This is not a concern to dischargers who closely follow the Super Order, because its requirement to take all samples within thirty days is designed to eliminate the effects of seasonality.Analyzing for seasonality can start by graphing the data by season and then running a two-way analysis of variance in which the well and season are the two factors in the model. If seasonal patterns are present (or significant), then the data should be adjusted to remove the seasonal pattern, prior to analysis. If insufficient data are available, when analysis of the data is initiated, then analysis may proceed as if seasonality is absent, until such time as a statistical determination of the seasonality in the data is possible. **IDT** has informed us that seasonality testing and seasonal adjustments are available in Version 1.3 of **GSAS**, and that these adjustments will be automatic, starting in Version 4.1; however, these features have not been reviewed.

Trends In Background: Again, if the proposed analysis is based on data collected over a range of dates, or if comparisons are being made with a concentration limit that is based on historical data, then the data need to be analyzed statistically to determine if the background concentration data exhibit trends. [Some methods are available within **GSAS** for doing this (e.g., Sen slopes) but, depending on the form of the data (for example, whether normally distributed or not), other methods such as regression may be preferable]. If background trends are present, then the data may need to be "detrended" (adjusted) prior to analysis, to take this into account.This issue is not of concern to dischargers using the Super Order's approach of relying upon concurrent background samples, because this practice eliminates the effects of trends in most cases.

Appropriateness Of Control Sites: Any inter-well comparison method [i.e., comparing data from downgradient (upgradient) wells with data from background wells] presumes that, with the exception of the possible effect of the facility, the control (upgradient) and downgradient sites should produce comparable measurements. The approach used under the Super Order assumes that the background wells represent the quality that ground water at the downgradient wells would show if there were no landfill. Rather than a statistical issue, this is a hydrogeologic issue that can have an effect upon the validity of statistical analysis. If either the discharger or the RWQCB finds that the facility's existing background wells cannot meet this performance standard, then the discharger should make suitable changes in the well network (e.g., install side-gradient wells to represent background water quality).

Independence Of Samples: One assumption underlying any of these statistical methods is that data collected at different sites or at the same site on different dates will be statistically independent. The choice between an analysis based on prediction intervals and one based on analysis of variance (ANOVA) will generally depend on an assessment of whether independent observations can be obtained over a short enough time period that the data can be expected to be equally affected by any potential discharge. This, again, is a hydrogeological issue rather than a statistical one, but it must be considered when an analysis method is proposed.The approach under the Super Order is to take all samples (downgradient and background) within thirty days. This requirement is intended to assure that all samples taken from a downgradient well either will show the presence of a release or will not. Therefore, for dischargers in strict compliance with the Super Order, the post-sampling purge provides reasonable sample independence.

USAGE OF GSAS, INCLUDING REPORTING PRACTICES

GSAS provides its output in many different forms and, therefore, care must be taken to report all the information that is necessary for the interpretation of the analysis. The following information should be included.

Outliers: A comprehensive quality control program is a necessary component to any monitoring program. In addition to such components as trip blanks, field blanks, and duplicate samples, the data still need to be scrutinized to determine whether any unreasonable values of unknown origin have been obtained. This is particularly important in the background data. **GSAS** provides some tools for this type of analysis (such as Box and Whisker plots), but because they are not part of the main stream of analysis, they must be requested. It is important that the data be scrutinized in some way for the presence of possible outliers. We have been in informed that outlier testing will be included in Version 3.3 of the **GSAS**, but we have not reviewed that feature.

Post Hoc Comparisons: **GSAS** provides appropriate analyses based on ANOVA or prediction intervals (to name two methods), but under California regulations, individual wells **must** be tested against background, whether the overall test for differences is significant or not. Thus, in reporting the results of any such analysis, the contrast test results must be reported.

Sample Size Considerations: The **GSAS** menus for analyses based on either ANOVA or prediction interval methods will select an appropriate method of analysis, whether it is based on the raw data, log-transformed data, or nonparametric methods. The sampling requirements will generally depend on whether the method used is parametric or nonparametric, but the increased number of water samples required at each well for a nonparametric analysis will be enforced only during the sampling period **following** the initial determination that nonparametric analysis is appropriate.For this reason, **reports to the RWQCB should indicate whether nonparametric methods were recommended during either the current sampling period or a previous sampling period**. In a sampling period for which the sample size was chosen in anticipation of a parametric analysis, but nonparametric methods were found to be required, the results of the

parametric analysis should be reported, despite the fact that the parametric method is no longer ideal. This procedure allows for a graceful transition to a method requiring a larger sample sizeCthe discharger collects the larger number of water samples required by the new statistical method during the following reporting period.In no instances should analysis be based solely on a method for which the current sample sizes are insufficient. In such a case, the choice between the parametric and nonparametric analysis should be made in consultation with a professional statistician; in the meantime, if the discharger wishes to immediately include the nonparametric analysis, then both results (i.e., from both the parametric and nonparametric) should be reported to the RWQCB.

Distributional Considerations: Just as the need for a nonparametric analysis may be apparent in some sampling periods but not in others, the need to transform the data may vary as well. Any changes of this sort [e.g., from the analysis of raw data, in one reporting period, to the analysis of log-transformed data in the next reporting period (or the other way around)] must be included in reports to the RWQCB.

ANALYSES NOT INCLUDED IN GSAS VERSION 1.3.

There are a number of statistical methods that may be useful approaches for a facility to use, but which are not implemented in Version 1.3 of **GSAS**. Two of these are noteworthy.

Discrete Retest Methods: For a large facility, it is possible that the large number of wells and constituents being tested in a given sampling period will add up to an unacceptably large probability of a false detection error. One approach to this problem is the use of a discrete retest, as described in 2550.7(e)(8)(E); this features the use of an adjusted error level (") for both the initial test and the two retest sample groups. This method is not fully implemented in **GSAS**, Version 1.3, since it does not allow the " level of individual comparisons to be adjusted to take into account the numbers of wells and constituents. It is our understanding that this feature has been fully implemented in Version 3.3 of the package, although this version has not been submitted for review.

Prediction Intervals Across Wells: The implementation of prediction intervals in **GSAS**, Version 1.3, uses it as a comparison method between background and a single downgradient well. It may be appropriate at times to compare the background to observations at a series of down-gradient wells.

GSAS Version 1.3 SUMMARY

Overall, we find **GSAS** to be a useful package for handling many of the more tedious aspects of the analysis of groundwater data using a user interface that is easy to follow, and we consider it to be applicable to dischargers who own or operate an MSW landfill subject to the Super Order, subject to the limitations described above. Software can not replace the care that must be exerted in the initial choice of a method of analysis; however, the MSW landfill owner or operator may find **GSAS** helpful in implementing the generalized statistical package adopted in the Super Order. Furthermore, **GSAS** contains a number of methods that may prove useful, subject to site-specific review and

acceptance, to dischargers wishing to deviate from the generic data analysis approach included in the Super Order.

----------<*>----------