TECHNICAL PAPER: DATA ANALYSIS FOR SOLID WASTE FACILITIES

.

The purpose of this document is to address common questions for the solid waste staff pertaining to the statistical analysis of groundwater samples at solid waste facilities. This document should be used in conjunction with the groundwater monitoring and sampling analysis plan. The statistical methods covered in this document include the most common statistical analyses used for groundwater monitoring samples at solid waste sites. For additional details, please refer to the EPA guidance documents listed on page 7 of this document.

INTRODUCTION

Statistical analysis of the groundwater data presented in monitoring reports for submission to the Virginia Department of Environmental Quality (VADEQ), Division of Waste Coordination should address the following:

A. Design of experiment
B. Outliers
C. Missing data
D. Evaluation of data below detection limits or quantitation limits
E. Checking assumptions (distributions, homogeneity of variances)
F. Selection of statistical method
G. Verification sampling strategy

A. DESIGN OF EXPERIMENT

The results of the statistical analysis can tell you only what the experiment was designed to explain. For example, up-gradient to down-gradient statistical comparisons will indicate if groundwater concentrations for a particular constituent are different up gradient of the landfill compared to down gradient of the landfill. This difference could be due to the landfill or due to natural site conditions. The facility must ensure that the design of the monitoring network and statistical experiment are designed to be able to detect a release of solid waste constituents from the landfill.

The facility should address natural spatial variation of groundwater constituents at a site when designing the monitoring network and type of statistical comparisons which will be performed. Two acceptable ways of dealing with spatial variability are to perform intra-well statistical comparisons only or to install additional up gradient or side gradient wells to account for natural variations at the site. If the facility possesses reliable pre-waste data (which have not been impacted by site activities) or can adequately demonstrate that inorganic constituent concentrations in wells which

are located down gradient from the landfill have not been impacted by site activities, the facility may petition the VADEQ for a variance from inter-well statistical comparisons. The variance petition should be written in accordance with 9 VAC 20-80-750 (Virginia Solid Waste Management Regulations (VSWMR)) and include hydro-geologic information about the site, a demonstration that inorganic constituent concentrations in down gradient wells have not been impacted by the landfill, information regarding the date waste was originally placed in the landfill, and the best estimate possible of groundwater flow at the site. If the facility is an older site, or it cannot be determined that inorganic constituent concentrations in groundwater from wells located down gradient of the landfill are not impacted by the landfill activities, the facility can install additional up gradient (or side gradient) wells to attempt to get a better estimate of natural variation at the site. Please note that the location of the additional up gradient or side gradient wells must be approved by VADEQ permitting staff.

The facility should also determine the number of background samples which will be necessary for the planned statistical analysis method and ensure that an adequate number of samples have been collected prior to the statistical comparisons required by the VSWMR. The facility should collect an adequate number of background dataset for inter-well statistical comparisons within one year, and an adequate number of background samples for intra-well statistical comparisons within two years. Background for inter-well statistical comparisons can be updated with each sampling event, unless there is an indication that background wells have been impacted by the landfill. Background for intra-well statistical comparisons can be updated every two years, unless there is indication of a release in the down gradient well. Please note that for intra-well comparison a two-year time window should be left between background for intra-well comparisons and compliance samples to ensure that samples associated with a slow release are not included in the background dataset.

The facility must sample for all constituents required by the VSWMR, unless it has been specified in the permit or a variance granted by the VADEQ that a facility may sample for constituents other than the full list required by the VSWMR.

B. OUTLIERS

Inconsistently large or small values (outliers) can be observed due to errors from sampling, laboratory, transportation, transcription, or actual extreme values. The historical background dataset should be screened for each well and constituent for the existence of outliers (USEPA 1992, section 6.2) using the method described by Dixon (1953) or another method approved by the VADEQ. Background observations, which are considered to be outliers, should not be included in the statistical analysis to preserve the power of the test to detect a release from the facility. If an extreme value occurs in compliance well during the compliance sampling event, the facility should collect a re-sample within the compliance period of the initial sample. This will enable the VADEQ to distinguish between an extreme value in a compliance well and an indication of a release from the facility. Background observations should be

evaluated to determine if data is normally distributed prior to running the outlier test.

C. MISSING DATA

If a sampling event results in a missing data value, an attempt to re-sample for the missing value should be made within the compliance period of the initial sampling event. It is recommended that the re-sample be collected as close to the initial sampling event as possible to minimize the effects of variation due to the differences in sample collection time and to allow additional time for a verification sample if one is needed.

D. DATA BELOW DETECTION LIMITS

The facility should use laboratory derived limits of detection and quantitation in the statistical analyses of groundwater data, as opposed to the detection and quantitation limits which have been published for a particular analytical method.

For data where the percentage of data below the laboratory limit of detection or laboratory limit of quantitation is less than 25 percent, the facility should replace the non-detects or non-quantified values with half the laboratory limit of detection or quantitation. However, when the percentage of non-detects or non-quantified values is greater than 25 percent and less than 50 percent, the mean and standard deviation should be adjusted using either Aitchison's adjustment (USEPA 1992 section 2.2.2 and Aitchison, 1955) or Cohen's adjustment (USEPA 1989 section 8.1.3 and Cohen, 1961). Extensive tables and computational details for Cohen's adjustment are also provided in Gibbons, 1994a. The approach for selection between the two methods is described in USEPA (1992) section 2.2.1.

E. CHECKING ASSUMPTIONS ASSOCIATED WITH THE TEST METHOD

Parametric statistical test methods assume that the data follow a certain distribution, for groundwater statistics the distributions usually are the normal and the log-normal distributions. The facility must verify that the distributional assumptions of a particular test method are valid prior to applying the statistical test method.

No testing of normality is needed when the percentage of non-detects or non-quantified values is greater than 50%, since a non-parametric statistical test method should be applied. Most parametric statistical tests for environmental data will assume the data are normally or log-normally distributed. The Shapiro-Wilk test, multiple group Shapiro-Wilk test or Filliben's correlation coefficient test should be applied to the dataset to determine the distributional form. To test for log-normality, the natural logarithms of the original data should be taken and tested for normality. The facility may use any other appropriate method for testing the distributional assumptions with approval by the VADEQ.

When the detection frequency is less than 50% or transformation fails to bring about normality, a non-parametric method should be used.

Non-parametric two- or multi-sample comparisons, such as the Wilcoxon rank sum test or the Kruskal-Wallis test assume that the dispersion for each group in the comparison is similar. This can be checked by comparing boxplots of each group.

F. SELECTION OF STATISTICAL METHOD

The facility should apply an appropriate statistical method consistent with the Virginia Solid Waste Management Regulations, 9 VAC 20-80-300.D.

Two- or Multi- Sample Comparisons

If a facility chooses to perform statistical comparisons using a two- or multi-way statistical test method (i.e. t-test, ANOVA, Wilcoxon rank sum, Kruskal-Wallis), the facility will need to collect a minimum of four samples per compliance period. As specified in the VSWMR the level of significance when performing these tests for individual well comparison shall be no less than 0.01 and no less than 0.05 for multiple comparisons. Due to the number of samples which need to be collected per compliance period most facilities prefer to apply the interval methods for statistical analysis associated with a compliance sampling event. However, when the intent of the statistical analysis is to show that mean/median concentration levels are similar between the background and compliance area (i.e. a first determination for an industrial or CDD landfill) the two- or multi-sample comparison statistical methods can be useful.

The facility should check distributional assumptions for both background and compliance datasets and check assumptions of homogeneity of variances prior to applying these tests.
The ANOVA test assumes data are normally or log-normally distributed and variances are homogeneous across groups. The CABF and Welch's t-tests assume data are normally or log-normally distributed and variances don't differ dramatically across groups (these tests account for some differences between variances). The Wilcoxon rank sum and Kruskal-Wallis tests assume that the distributions of the two groups are similar (though undetermined).

Interval Method

Statistical interval methods commonly applied in groundwater data analysis are the confidence interval, prediction interval, and tolerance interval. Prediction and tolerance intervals are often applied for compliance sampling events in Detection, Assessment, Phase I, and Phase II monitoring programs and for establishing background-based, groundwater protection standards, since only one initial sample per well is required during the compliance period. Confidence intervals are often applied for comparisons to a groundwater protection standard which is based on a mean or median value.

For all interval methods, the facility should check the normality or log-normality of the background dataset and the percentage of non-detects in the background dataset. If the background dataset is normally or log-normally distributed, and there are less than 50% non-detects, then a parametric interval can be calculated. If a distribution cannot be established for the background dataset or 50% or more of the data are non-detects, the facility should apply a non-parametric statistical limit.

Suggested sample sizes for the parametric and non-parametric versions of the above interval methods are provided in the attached table. Please note that these methods can lead to a higher false positive rate or lower statistical power with a smaller sample size. However, a statistical analysis can be conducted with a smaller dataset than the suggested size at any time.

It is the responsibility of the facility to collect an adequate number of background samples for the proposed statistical interval methods prior to the statistical analysis event required by the VSWMR. False positive and false negative rates associated with confidence, prediction and tolerance intervals must be protective of human health and the environment. If the facility chooses to apply a false positive rate of less than .01, the facility must include in the report a demonstration that a lower false positive rate will provide adequate statistical power to detect a release from the facility. Adequate statistical power is the ability to detect a three standard deviation increase above the mean with 50% power and a four standard deviation increase above the mean with 80% power.

Control Charts

The Shewhart-CUSUM control chart can be applied as an intra-well statistical test method. Please note that a variance from inter-well statistical comparisons must be granted by the VADEQ prior to applying an intra-well only monitoring program. Details of how to apply Shewhart-CUSUM control charts can be found in EPA 1992 (section 7). Please note that the background dataset can be updated every two years if there is no indication of an impact from the facility (increasing trend or significant result). The facility should leave a two-year time window between the background dataset and the compliance event to ensure that data associated with a slow release from the facility are not incorporated into the background dataset.

Other Methods

In the event the facility has selected any other method listed in the Virginia Solid Waste Management Regulations, the facility will collect the appropriate number of samples and shall maintain an appropriate level of significance mentioned above. If the facility prefers to apply a statistical method that is not in listed in the VSWMR, the facility must receive approval from the VADEQ prior to applying the test method.

Comparison of Compliance Well Data To A Standard During Assessment Or Corrective Action Monitoring

In accordance with sections 9 VAC 20-80-300.B.3 and 300.C.4 (VSWMR) the compliance data shall be compared to the groundwater protection standard (GWPS) if down gradient well concentrations exceed established background concentrations for Table 5.1 constituents. If a maximum contaminant level (MCL) is promulgated or alternate concentration limit (ACL) is established for a constituent, and the ACL or MCL is greater than the background limit (or statistically determined background level), the ACL or MCL is the ground-water protection standard. All new concentrations in the assessment or corrective action wells should be compared to the standard (i.e., ACL or MCL) using the lower normal confidence limit computed from at least four sampling values collected during the compliance period. The level of confidence of the interval should be 80% for a sample size of 4-7, and 90% for a sample size of 8-10 to ensure that the comparison has adequate power to detect an exceedance above the groundwater protection standard.

If the groundwater protection standard for a constituent is based on background data and exceeds the MCL or ACL, then the individual point of compliance

measurements will be compared to the background limit and not the MCL or ACL.

However, for a particular sampling event, if the established groundwater protection standard is less than the VADEQ accepted quantitation limit (QL) then the QL becomes the standard for that sampling event, and the compliance well data will be compared to the QL.

G. VERIFICATION SAMPLING

The principal advantage of taking a verification sample is to maintain an acceptable site-wide false positive rate while the statistical test has adequate power to detect a release from the facility if it occurs. A verification sampling strategy involves collection of a pre-planned number of additional samples. A facility may choose to apply verification samples as follows:
The 1-of-m approach was initially suggested by Davis and McNichols (1987). The facility can take as many as m samples during the compliance period of the initial sampling event and if the 1-of-m (usually m=1 to 3) sample is below a prediction or tolerance limit, the constituent is said to have "passed" the test at that well. If the facility chooses to apply the verification sampling strategy, the alpha value should be modified as following:

a. Select a default value for α = 0.01
$$\alpha = 0.01$$
b. Pass the first or one of one verification resamples, adjust alpha
$$\alpha = (1-.95^{\frac{1}{k}})^{\frac{1}{2}}$$
c. Pass the first or one of two verification resamples, adjust alpha
$$\alpha = (1-.95^{\frac{1}{k}})^{\frac{1}{3}}$$
d. Pass the first or two of two verification resamples, adjust alpha
$$\alpha = \sqrt{1-0.95^{\frac{1}{k}}}\sqrt{\frac{1}{2}}$$

Where k is the number of comparisons and α is the site-wide false positive rate. Please note that alpha can not be less than 0.01 unless the facility shows that the statistical comparison has at least as much statistical power as the EPA reference power curves (EPA 1992, Appendix B). Since the verification sampling is pre-planned, the facility can adjust the upper statistical limit calculated for background to account for the fact that the verification samples will be collected. Please note that the regulations do not allow a facility to disregard the statistical evaluation in a situation when the facility is unable to collect a verification sample. Therefore, if the facility would like to take a verification sample, it should be taken during the compliance period of the initial sampling event and the statistical result must include the verification sample prior to submitting it to the VADEQ. The verification sample must be independent from the initial sample.

For questions or comments, please contact:
Hasan Keceli
Statistician
Office of Waste Programs
hasan.keceli@deq.virginia.gov
(804) 698-4246

REFERENCES

Aitchison, J. On the distribution of a positive random variable having discrete

probability mass at the origin, Journal of American Statistical Association, 50(272), 901-908 (1955).

Cohen, A.C. Tables for maximum likelihood estimates: Singly truncated and singly censored samples, Technometrics, 3, 535-541 (1961).

Davis, C. B., and McNichols, R. J. (1987). One-side intervals for at least p of m observations from a normal population on each of r future occasions. Technometrics, 29 359-370.

Dixon, W. J. Processing data for outliers, Biometrics 9 , 74-89 (1953)

Gibbons, R.D. Statistical Methods for Groundwater Monitoring, John Wiley and Sons, Inc., 1994.

Gibbons, R.D. Some conceptual and statistical issues in analysis of groundwater monitoring data, Environmetrics, 7, 185-199 (1996).

USEPA, Statistical analysis of groundwater monitoring data at RCRA facilities. Addendum to Interim Final Guidance. Office of Solid Waste, July 1992.

USEPA, Statistical analysis of groundwater monitoring data at RCRA facilities - Interim Final guidance (April 1989).

VADEQ, Virginia Solid Waste Management Regulations, Department of Waste Management (March 1993).

Approved by: Sanjay V. Thampi

Director
Office of Waste Technical Suopport

Date: 3/25/08

Note: This doccument has been reviewed by Waste Division staff.

TABLE 1
SUGGESTED MINIMUM BACKGROUND SAMPLES

	Parametric	Non-parametric	Non-parametric Interval % Confidence
CABF/Welch's T-test	4	NA	NA
Wilcoxon Rank Sum	NA	5	NA
Confidence Interval	4	NA	NA
Tolerance Interval	8	19	95%
Prediction Interval	8	13	99%#
Shewhart CUSUM Chart+	8	NA	NA

* The above tests can be used with fewer samples, however it will increase the false positive rate.
# Includes one verification re-sample, use 19 samples for a 95% Prediction Interval with no verification resamples.
+ For Intra-well testing only.
NA Not Applicable.