GUIDELINE 2 - STATISTICAL ANALYSIS OF GROUNDWATER MONITORING DATA FROM SOLID WASTE MANAGEMENT FACILITIES

North Dakota Department of Health - Division of Waste Management
918 E. Divide Ave., 3rd Fl., Bismarck, ND 58501-1947
Telephone: 701.328.5166 • Fax: 701.328.5200 • Website: www.ndhealth.gov/wm
Updated 04-2009

I. Introduction

Statistical analysis methods are applicable to all existing units, new units, and lateral expansions of existing units that are required to conduct groundwater monitoring. The use of statistical methods to evaluate monitoring data is necessary for the duration of the monitoring program inclusive of the postclosure period.

The owner or operator must indicate in groundwater monitoring strategies or plans the statistical method that will be used in the analysis of groundwater monitoring results (NDAC 33-20-13-02). The number of samples collected, and the frequency of collection, must be consistent with the statistical method selected.

Several options for analysis of groundwater data are provided. Other methods are not excluded if they can be shown to meet statistical performance standards. The recommended methods include both parametric and non-parametric procedures which differ primarily in constraints placed by the statistical distribution of the data. Control chart, tolerance interval, and prediction interval approaches may also be applied.

The owner or operator must conduct the statistical comparisons between upgradient and downgradient wells after completion of each sampling event and receipt of validated data (NDAC 33-20-13-02). The Department recommends that the statistical procedure conforms to the performance standard of a Type I error level of no less than 0.10 for inter-well comparisons. Control chart, tolerance interval and prediction interval approaches must incorporate decision values which are protective of human health and the environment. Generally, this is meant to include a significance level of at least 0.10. Procedures to treat data below analytical method detection levels and seasonality effects are necessary prior to statistical analysis.

This document has been prepared by the Department for the purpose of assisting owners and operators who conduct groundwater monitoring to fulfill regulatory and permit requirements. Questions and comments are welcome, and can be addressed to the Division of Waste Management.

II. Technical Considerations

In most cases, the Department recommends one of five statistical methods for evaluating groundwater monitoring data. Different methods may be selected for each groundwater quality constituent. The appropriateness of a method must be substantiated by demonstrating that the distribution of the data for that constituent is appropriate for the method. Selection of a specific method is described in the USEPA "Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities - Interim Final Guidance" (USEPA, 1989) and is also discussed in "Statistical

Analysis of Groundwater Monitoring Data at RCRA Facilities - Addendum to Interim Final Guidance" (USEPA, 1992). The methods include the following:

1. Parametric analysis of variance (ANOVA);
2. Rank based (non-parametric) ANOVA with multiple comparisons;
3. Tolerance prediction interval;
4. Control chart; and
5. An alternative statistical method.

These statistical analysis methods are necessary to determine whether a significant increase over background has occurred. The statistical analysis of monitoring data occurs after receiving validated results from each sampling and analysis event. If an alternative method is used, then the Department must be notified and a justification for its use must be provided.

Performance standards are necessary for a statistical analysis method to provide reliable and valid conclusions. These standards are:

1. Applicability to actual distribution of the data;

2. Individual well comparisons to background groundwater quality or a groundwater protection standard shall be done at a Type I error level no less than 0.10 or, if the multiple comparisons procedure is used, the experiment-wise error rate shall be no less than 0.10;

3. If a control chart is used, the type of chart and associated parameter values shall be protective of human health and the environment;

4. The level of confidence and percentage of the population contained in an interval shall be protective of human health and the environment;

5. Account for data below the limit of detection (less than PQL) in a manner that is protective of human health and the environment; and, if necessary,

6. Account for seasonal and spatial variability and temporal correlation of the data.

   The statistical performance standards provide a means to limit the possibility of making false conclusions from the monitoring data. The specified error level of 0.10 for individual well comparisons for probability of Type I error (indication of contamination when it is not present, or false positive) essentially means that the analysis is predicting with 90 percent confidence that no significant increase in contaminant levels is evident. The corollary is that there is only a 10 percent chance that a Type II error (failure to detect a significant increase in constituent concentration, or false negative) has occurred.

   Nondetected results must be treated in an appropriate manner or their influence on the statistical method may invalidate the statistical conclusion. Further discussion of nondetected results is found later in this section. In addition to the statistical guidance provided by USEPA (1989), the following references may be useful for selecting other methods (Dixon and Massey, 1969; Gibbons, 1976; Aitchison and Brown, 1957; and Gilbert, 1987).

III. Multiple Well Comparisons

The goal of monitoring well placement is to provide representative samples of water quality both upgradient and downgradient of the waste management unit. In this manner, changes in the quality of groundwater can be determined. If more than two wells (upgradient and downgradient combined) are screened in the same stratigraphic unit then the appropriate comparison method is a multiple well comparison using analysis of variance or ANOVA. The analysis of variance test compares the average concentration among wells to determine if they are from the same continuous distribution. The ANOVA test includes both parametric and non-parametric procedures. Chemical data from groundwater tends to follow a log normal distribution (USEPA, 1989) and data usually needs to be transformed prior to applying a parametric ANOVA procedure.

By making a log transformation, data will generally be converted to a normal distribution. By applying a chi-squared procedure, probability plots or other normality tests on the residuals (errors) from the ANOVA procedure, the normality of the transformed data can be determined. In addition, the variance of data from each well in the comparison must be approximately equivalent; this condition can be checked using Bartlett's test. Both Bartlett's test and four normality tests are presented in USEPA (1989).

If the transformed test data does not conform to the normality assumption, a non-parametric ANOVA procedure should be used. The non-parametric statistical procedures are not dependent on the mathematical properties of a specified distribution. The non-parametric equivalent to the parametric ANOVA is called the Kruskall-Wallis Test which uses ranking methods to compare the data.

If the data displays seasonality (consistent time dependent increases or decreases in parameter values), a two-way ANOVA procedure should be used. If the seasonality can be corrected, a one-way ANOVA procedure may still be appropriate. Methods to treat seasonality are described in USEPA (1989).

ANOVA procedures determine whether different wells have significantly different concentrations of constituents. ANOVA procedures are followed by multiple comparisons procedures that are used to discriminate between wells showing significant differences. The multiple comparison procedures test the contrast between the mean value of each monitoring parameter for the background to the compliance boundary well. For limitations on multiple comparison tests when more than five wells are used, refer to USEPA (1989).

If data lacks variability between upgradient and downgradient wells, other methods, including tolerance intervals and prediction intervals, may be used in place of analysis of variance. These methods may be especially appropriate where the site geology is very homogeneous. Both methods are similar and consist of constructing confidence intervals about the mean of the background well data. The confidence interval should be one-sided since the objective of monitoring is to detect significant changes that are greater than the background mean. The confidence level should set to 90 percent. These methods are discussed also in USEPA (1989).

IV. Individual Well Comparisons

Sites with complex geology may have stratigraphic units which are monitored by only one background and one downgradient well. Under these conditions the ANOVA procedure, which requires a minimum of three wells, may not be used. In these instances, other statistical methods would be required to compare background and compliance boundary data.

A number of statistical procedures can be found in statistic references, for the comparison of means where variances are equal or where variances are unequal. The distribution assumptions of the method selected must be satisfied by the distribution of the data, and the confidence level should be no less than 90 percent for a Type I error.

V. Intra-Well Comparisons

Intra-well comparisons, where data of one well are evaluated over time, are useful in evaluating trends in individual wells and for identifying seasonal effects in the data. During initial rounds of groundwater monitoring, seasonal effects should not be misconstrued for either positive or negative trends. The intra-well comparison methods do not provide evaluations of background data and compliance boundary data.

Some existing facilities may not have valid background data. Intra-well comparisons may represent the only valid comparison available. For example, a facility constructed on a hilltop with radial groundwater flow around the facility, may not have an upgradient well. If the geology is not similar nearby, it may not be able to clearly establish background groundwater quality from adjacent wells.

A significant positive trend in the data from a downgradient monitoring well may provide sufficient evidence to conclude that the landfill unit is affecting the well. Such trends should be evident in several parameters since contaminant migration from a landfill will result in an influx of multiple constituents into the groundwater flow system.

Control charts may be used for intra-well comparisons but are only appropriate for uncontaminated wells. If a well is intercepting a release, then it is already in an out-of-control state which violates the principal assumption underlying control chart procedures.

Time series analysis (plotting concentrations over time) is extremely useful for identifying trends in monitoring data. Such data may be adjusted for seasonality effects to aid in assessing the degree of change over time. Guidance for intra-well comparison techniques and limitations of the techniques are provided in USEPA (1989).

VI. Treatment of Nondetections

The treatment of data below the detection limit of the laboratory's analytical method is dependent upon the number or percentage of nondetections (NDs) and the statistical method employed. If a large proportion (greater than 50 percent) are present, a professional statistician should be consulted. USEPA (1989) provides general guidance on this topic which is briefly summarized here.

If the amount of data below detection is less than 15 percent, this data may be substituted with values of one half of the reported method detection limit (MDL/2) or one half the method's practical quantitation limit (PQL/2). If more than 15 percent of the data are NDs then non-parametric test procedures are recommended for evaluating the monitoring data. The non-parametric methods will treat the NDs values as ties; this procedure is less sensitive to error than parametric methods.

For NDs which represent greater than 50 percent of the data set, the test of proportions should be used. The test of proportions identifies the number of detections in a compliance well, relative to the background well, which is an indication of a statistically significant increase. These and other methods of handling nondetection data are discussed in USEPA (1989).

VII. Comparisons to Regulatory Limits

If groundwater data must be evaluated to determine whether a constituent has exceeded a regulatory limit (e.g., an MCL), a confidence interval approach based on the distribution of the data should be used. The confidence interval is designed to contain the true mean of the data with a specified level of confidence (generally 90 percent at the lower limit). The lower limit is then compared to the regulatory value and if the lower limit is larger, it is considered evidence that the regulatory level has been exceeded.

The data or log transformations of the raw data must meet normality assumptions when the student t-distribution is used. When nonparametric methods are employed, a minimum of seven independent data values must be available from a sampling event that is not affected by seasonality. Other distributions may be used to create the confidence interval so long as the data follows the assumed distribution. For further guidance, refer to USEPA (1989).

VIII. Statistical Analysis of Municipal Waste Landfills

Federal regulations, 40 CFR Parts 257 and 258, require owners or operators of municipal waste landfills to evaluate groundwater monitoring data using a statistical method provided in §258.53(g) that meets the performance standard of §258.53(h). §258.53(g) contains a provision allowing for an alternative statistical method as long as the performance standards of §258.53(h) are met. The following excerpt from 40 CFR Part 258(g) - (I) addresses statistical analysis at municipal waste landfills:

"(g) The owner or operator must specify in the operating record one of the following statistical methods to be used in evaluating groundwater monitoring data for each hazardous constituent. The statistical test chosen shall be conducted separately for each hazardous constituent in each well.

(1) A parametric analysis of variance (ANOVA) followed by multiple comparisons procedures to identify statistically significant evidence of contamination. The method must include estimation and testing of the contrasts between each compliance well's mean and the background mean levels for each constituent.

(2) An analysis of variance (ANOVA) based on ranks followed by multiple comparisons procedures to identify statistically significant evidence of contamination. The method must include estimation and testing of the contrasts

between each compliance well's median and the background median levels for each constituent.

(3) A tolerance or prediction interval procedure in which an interval for each constituent is established from the distribution of the background data, and the level of each constituent in each compliance well is compared to the upper tolerance or prediction limit.

(4) A control chart approach that gives control limits for each constituent.

(5) Another statistical test method that meets the performance standards of §258.53(h). The owner or operator must place a justification for this alternative in the operating record and notify the State Director of the use of this alternative test. The justification must demonstrate that the alternative method meets the performance standards of §258.53(h).

(h) Any statistical method chosen under §258.53(g) shall comply with the following performance standards, as appropriate:

(1) The statistical method used to evaluate groundwater monitoring data shall be appropriate for the distribution of chemical parameters or hazardous constituents. If the distribution of the chemical parameters or hazardous constituents is shown by the owner or operator to be inappropriate for a normal theory test, then the data should be transformed or a distribution-free theory test should be used. If the distributions for the constituents differ, more than one statistical method may be needed.

(2) If an individual well comparison procedure is used to compare an individual compliance well constituent concentration with background constituent concentrations or a groundwater protection standard, the test shall be done at a Type I error level no less than 0.01 for each testing period. If a multiple comparisons procedure is used, the Type I experiment wise error rate for each testing period shall be no less than 0.05; however, the Type I error of no less than 0.01 for individual well comparisons must be maintained. This performance standard does not apply to tolerance intervals, prediction intervals, or control charts.

(3) If a control chart approach is used to evaluate groundwater monitoring data, the specific type of control chart and its associated parameter values shall be protective of human health and the environment. The parameters shall be determined after considering the number of samples in the background data base, the data distribution, and the range of the concentration values for each constituent of concern.

(4) If a tolerance interval or a predictional interval is used to evaluate groundwater monitoring data, the levels of confidence and, for tolerance intervals, the percentage of the population that the interval must contain, shall be protective of human health and the environment. These parameters shall be determined after considering the number of samples in the background data base, the data

distribution, and the range of the concentration values for each constituent of concern.

(5) The statistical method shall account for data below the limit of detection with one or more statistical procedures that are protective of human health and the environment. Any practical quantitation limit (PQL) that is used in the statistical method shall be the lowest concentration level that can be reliably achieved within specified limits of precision and accuracy during routine laboratory operating conditions that are available to the facility.

(6) If necessary, the statistical method shall include procedures to control or correct for seasonal and spatial variability as well as temporal correlation in the data.

(I) The owner or operator must determine whether or not there is a statistically significant increase over background values for each parameter or constituent required in the particular groundwater monitoring program that applies to the MSWLF unit, as determined under §§258.54(a) or 258.55(a) of this part.

(1) In determining whether a statistically significant increase has occurred, the owner or operator must compare the groundwater quality of each parameter or constituent at each monitoring well designated pursuant to §258.51(a)(2) to the background value of that constituent, according to the statistical procedures and performance standards specified under paragraphs (g) and (h) of this section.

(2) Within a reasonable period of time after completing sampling and analysis, the owner or operator must determine whether there has been a statistically significant increase over background at each monitoring well."

IX. References

Aitchison, J., and J.A.C. Brown, (1969). "The Lognormal Distribution; "Cambridge University Press; Cambridge.

Dixon, W.J. and F.J. Massey, Jr., (1969). "Introduction to Statistical Analysis;" 3rd Edition; McGraw-Hill Book Co.; New York, New York.

Gibbons, J.D., (1976). "Non-parametric Methods for Quantitative Analysis;" Holt, Rinehart, and Winston Publishing Co.; New York, New York.

Gilbert, R.O., (1987). Statistical Methods for Environmental Pollution Monitoring; Van Nostrand Reinhold Co.; New York, New York.

USEPA, (1989). "Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities - Interim Final Guidance;" EPA/530-SW-89-026; U.S. EPA; Office of Solid Waste Management; Washington, D.C.