Statistical Analysis Plan

MERRIMACK STATION COAL ASH LANDFILL

Bow, New Hampshire

Sanborn, Head & Associates, Inc.

Digitally signed by: Sanborn, Head & Associates, Inc.
DN: CN = Sanborn, Head & Associates, Inc. email = lalic@sanbornhead.com C = US O = Sanborn, Head & Associates, Inc.
Date: 2024.01.23 16:36:54 -05'00'

Prepared for GSP Merrimack LLC
File No. 2025.14
January 23, 2024

TABLE OF CONTENTS

1.0 INTRODUCTION

Groundwater monitoring at the Merrimack Station Coal Ash Landfill site (Site) in Bow, New Hampshire is required pursuant to 40 CFR Part 257.90. Sanborn, Head & Associates, Inc. (Sanborn Head) prepared this Statistical Analysis Plan for detection monitoring at the Site required by 40 CFR Part 257.94. The statistical method and professional certification requirements are specified in 40 CFR Part 257.93. A qualified professional engineer certification is provided in Appendix A. This Statistical Analysis Plan and the services provided by Sanborn Head are subject to the Limitations provided in Appendix B.

2.0 OVERVEIW AND APPLICABLE REFERENCES

The methods described in this Statistical Analysis Plan are selected for evaluating the groundwater monitoring data for detection monitoring at the Site, as required by 40 CFR Parts 257.93 and 257.94. Pursuant to 40 CFR Part 257.93(f), the following statistical methods may be used for detection monitoring:

1. Parametric analysis of variance with multiple comparison procedures;
2. Analysis of variance with multiple comparison procedures;
3. Tolerance or prediction interval procedures;
4. Control charts; and
5. Other statistical test method that meets performance standards of CFR Part 257.93(g).

Additional statistical references used in the development of this Statistical Analysis Plan are the U.S. Environmental Protection Agency Unified Guidance (USEPA, 2009), Helsel (2012), and Helsel et al. (2020).

The proposed prediction limit approach for detection monitoring is based on a site-wide false positive rate of 10% over a year of testing and adequate statistical power, as recommended in the USEPA Unified Guidance. Unless noted otherwise, supporting statistical tests for parametric fits, trends, outliers, and other procedures will use a significance level of 0.01.

This Statistical Analysis Plan assumes that suitable alternatives to the named supporting statistical tests for parametric fits, trends, outliers, and other procedures may be used, and the use of suitable alternatives is consistent with this plan. If a substantive change in process or detection monitoring statistical method is made, then a revised Statistical Analysis Plan may be needed.

3.0 DETECTION MONITORING TESTS

For detection monitoring, compliance data are compared with background data using statistical tests. An initial eight background samples were collected at Site monitoring wells in 2016 and 2017. Background data are periodically updated after at least four new compliance data are collected. An intrawell background approach is used for the Site, where background data are based on historical data selected for each location-constituent pair. For example, the background data for sulfate at monitoring well SB-1 includes historical sulfate concentrations at SB-1.

If background data across multiple wells are not significantly different, then background data for multiple wells may be pooled to provide a larger background data set. Before pooling, differences in variance and distribution across multiple wells will be tested statistically. For example, the Fligner-Killeen test may be used for testing differences in variance, and the Kruskal-Wallis one-way analysis of variance may be used for testing differences in groups. These tests are relatively robust to varying distributions, and suitable alternative tests may be used.

The primary statistical method for detection monitoring are parametric or nonparametric prediction limit tests. One-tailed upper prediction limits will be used to test for increases relative to background for all analytes except pH. Two-tailed prediction limits will be used for pH to test for both increases and decreases relative to background. The parametric methods assume the data follow a known distribution, such as a normal or lognormal distribution, and the nonparametric methods do not assume a known distribution is followed. If a fitting distribution is identified, then typically the parametric methods are preferred.

- The closest parametric fit for a background dataset is selected using the "Ladder of Powers" approach described in Helsel et al. (2020) using probability plots or statistical criteria, such as the Shapiro-Francia test statistic.
- The closest parametric fit is then tested using Shapiro-Wilk, Shapiro-Francia, probability plot correlation coefficient, or similar tests.
- If the background dataset fits the parametric distribution, then a parametric prediction limit is calculated for future observations with 1-of-2 retesting strategy, semi-annual sampling, a Site-wide false positive rate of 10%, and an "acceptable" or "good" statistical power (USEPA, 2009; Section 19.3.1 and Appendix D Table 19-10).
- If data do not fit a parametric distribution, then a non-parametric prediction limit is calculated for future observations with 1-of-2 retesting strategy, semi-annual sampling, a Site-wide false positive rate of 10%, and an "acceptable" or "good" statistical power (USEPA, 2009; Section 19.4.1 and Appendix D Table 19-19).

If the background dataset includes non-detect data, then the following procedures will be used.

- If all the background data are non-detect, then the Double Quantification rule is used, where a "confirmed exceedance is registered if any well-constituent pair in the '100% non-detect' group exhibits quantified measurements (i.e., at or above the reporting limit [RL]) in two consecutive sample and resample events" (USEPA, 2009).
- If 50% or greater of the background data are non-detect, then a non-parametric prediction limit is used. This typically results in selecting the largest or second-largest detected value as the prediction limit.
- If less than 50% of the background data are non-detect, then Kaplan-Meier method, maximum likelihood estimation, or robust regression on order statistics is used to estimate summary statistics. These methods use the distribution of the detect and non-detect data to estimate mean and standard deviation of the data. The summary statistics are then used for the prediction limit test. If the background data do not fit a parametric distribution, then a non-parametric prediction limit is used.

4.0 UPDATING BACKGROUND DATA

Background data may be updated after at least four new, statistically independent samples are collected. The criterion of four samples is based on the suitability of statistical tests for comparison of the new data to existing background data (USEPA, 2009). Proposed background data will not include data that were statistically significant increases (SSIs) during detection monitoring unless those SSIs were resolved through an alternative source demonstration. The proposed background data will be tested for the following conditions.

- Similarity of proposed new background data to existing background data is tested using a rank-sum test, such as Mann-Whitney U test or Peto-Peto test, and a significance level of 0.05. If rank-sum tests are not applicable because the proportion of non-detect values in the data was too great, then similarity may be judged by comparing percent detections or other measures.
- Potential outliers are identified by reviewing visualizations, such as time series and quantile-quantile plots, of the combined new and existing background data. Potential outliers are then tested using either Dixon's Test for sample size less than 20 or Rosner's Tests for sample size of 20 or greater. Prior to testing, the data may be transformed using the best parametric fit. If potential outlier is determined to be not representative of background, then it is not included in the background dataset.
- Temporal trends, seasonality, and autocorrelation are screened for by reviewing timeseries plots. If patterns are identified visually, then methods described in the USEPA Unified Guidance are used to test for and correct for temporal variability. Prior to correcting for an increasing trend, either a significantly increasing trend must be present in both the downgradient compliance well and the upgradient well background data or there must be a determination that the increasing trend is not evidence of possible CCR impacts at the Site.

Three approaches are potentially used for handling a situation where proposed background data are identified as either significantly different from existing background, outliers, or having temporal variability that cannot be corrected for. First, if recent data are considered potentially not representative of current background conditions, then background data are not updated. Second, if recent data are considered more representative of current background conditions than older data, then older data may be dropped from background so that background data better reflect current conditions. Lastly, if both recent data and older data are considered representative of current background conditions, then background data may be updated and include both recent and older data. Discussion of processes for updating background is provided in the USEPA Unified Guidance, and alternative source demonstrations may be used as evidence that data reflect current background conditions.

5.0 REFERENCES

U.S. Environmental Protection Agency, 2009. *Statistical analysis of groundwater monitoring data at RCRA facilities: EPA 530-R-09-007*.

Helsel, D.R., 2012. *Statistics for censored environmental data using Minitab and R.* New York, John Wiley & Sons.

Helsel, D.R., Hirsch, R.M., Ryberg, K.R., Archfield, S.A., and Gilroy, E.J., 2020. *Statistical methods in water resources: U.S. Geological Survey Techniques and Methods, Book 4, Chapter A3*.

P:\2000s\2025.14\Source Files\202401 Stats Cert\202401 Stats Cert.docx

Appendix A

Qualified Professional Engineer Certification

Appendix A
Qualified Professional Engineer Certification

I certify that the information in this Statistical Method Selection Certification, dated January 23, 2024 (the "Statistical Analysis Plan"), is appropriate for evaluating the groundwater monitoring data for detection monitoring, as required by 40 CFR Part 257.93(f), subject to the assumptions and limitations contained within the Statistical Analysis Plan. The Statistical Analysis Plan was prepared by Sanborn, Head & Associates, Inc. for the Merrimack Station Coal Ash Landfill site located in Bow, New Hampshire.

Harrison R. Roakes
Printed Name of Licensed Professional Engineer

Signature

15920	New Hampshire	1/23/2024
License Number	Licensing State	Date

Appendix B

Limitations

Appendix B
Limitations

1. The conclusions and recommendations presented in this Statistical Analysis Plan are based in part on the data obtained from a limited number of samples from widely-spaced locations. The sample results indicate conditions only at the specific location and time. They do not necessarily reflect variations that may exist between or within such locations, and the nature and extent of variations between or within these locations may not become evident until further investigation or remediation is initiated. The validity of the conclusions is based in part on assumptions Sanborn Head has made about conditions at the site. If conditions different from those described become evident, then it will be necessary to reevaluate the conclusions of this Statistical Analysis Plan.

2. Sanborn Head relied upon data provided by analytical laboratories and did not independently evaluate the reliability of these data. Additionally, variations in the types and concentrations of analytes and variations in their distributions may occur due to the passage of time, water table fluctuations, precipitation and recharge events, and other factors.

3. The conclusions and recommendations contained in this report are based in part upon various types of chemical data, historical and hydrogeologic information developed during previous studies, and statistical method guidance and references. While Sanborn Head has reviewed those data and information as stated in this report, any of Sanborn Head's interpretations, conclusions, and recommendations that have relied on that information will be contingent on its validity. Should additional chemical data, historical information, hydrogeologic information, or reference material become available in the future, such information should be reviewed by Sanborn Head and the interpretations, conclusions, and recommendations presented herein should be modified accordingly.

4. This Statistical Analysis Plan was prepared for the exclusive use of GSP Merrimack LLC (GSP) for specific application for groundwater detection monitoring as required by 40 CFR Part 257.94 for the Merrimack Station Coal Ash Landfill site located in Bow, New Hampshire and was prepared in accordance with generally-accepted hydrogeologic practices. No warranty, express or implied, is made.