N61165.AR.002966
CNC CHARLESTON
5090.3a

LETTER TRANSMITTING COMPLETE RESOURCE CONSERVATION AND RECOVERY ACT FACILITY INVESTIGATION TECHNICAL MEMO TO REPLACE INCOMPLETE VERSION SENT 2 JUNE 1995 CNC CHARLESTON SC

6/8/1995

ENSAFE/ ALLEN AND HOSHALL

EnSafe / Allen & Hoshall
a joint venture for professional services

June 8, 1995

Program Management Office
Shelby Oaks Plaza
5909 Shelby Oaks Dr.
Suite 201
Memphis, TN 38134
Phone (901) 383-9115
Fax (901) 383-1743

EnSafe/Allen & Hoshall Branch Offices:

Charleston
935 Houston Northcutt Blvd.
Suite 113
Mt. Pleasant, SC 29464
Phone (803) 884-0029
Fax (803) 856-0107

Cincinnati
400 TechneCenter Dr.
Suite 301
Milford, OH 45150
Phone (513) 248-8449
Fax (513) 248-8447

Pensacola
2114 Airport Blvd.
Suite 1150
Pensacola, FL 32504
Phone (904) 479-4595
Fax (904) 479-9120

Norfolk
303 Butler Farm Road
Suite 113
Hampton, VA 23666
Phone (804) 766-9556
Fax (804) 766-9558

Raleigh
5540 Centerview Drive
Suite 205
Raleigh, NC 27606
Phone (919) 851-1886
Fax (919) 851-4043

Nashville
311 Plus Park Blvd.
Suite 130
Nashville, TN 37217
Phone (615) 399-8800
Fax (615) 399-7467

Dallas
4545 Fuller Drive
Suite 326
Irving, TX 75038
Phone (214) 791-3222
Fax (214) 791-0405

VIA HAND DELIVERY

Commanding Officer
Attn: Matthew A. Hunt, Code 1877
Southern Division, Naval Facilities Engineering Command
2155 Eagle Drive
Charleston, South Carolina 29411-0068

Re: NAVBASE Charleston RFI, Contract Number: N62467-89-D-0318,
Background Evaluation Technical Memo

Dear Mr. Hunt:

EnSafe/Allen & Hoshall (E/A&H) discovered that the technical memo submitted to you regarding background on June 2, 1995 was an incomplete version. A number of attachments that were referenced in the document were inadvertently left out of the submittal package. I have enclosed a revised memo which includes all of the attachments. Also, I will be forwarding copies of the revised version to Doyle Britton and Joe Bowers so that they can complete a review of the document. Concurrence from USEPA and SCDHEC on an acceptable means to calculate background remains a critical issue with the development of the Draft Zone H RFI Report. E/A&H feels strongly that, in the best interest of the Zone H report, a meeting with the regulatory agencies to discuss the issue is urgently needed since the development of the report is well underway.

I apologize for any inconvenience this may have caused and look forward to an opportunity for E/A&H to discuss this issue with the respective agencies soon. If you have any questions or if I can be of assistance please do not hesitate to call me at 884-0029.

Sincerely,
EnSafe/Allen & Hoshall
A Joint Venture in Professional Services

Todd Haverkost

By: Todd Haverkost
Task Order Manager

Attachment

cc: Doyle Britton, U.S.E.P.A.
Joe Bowers, SCDHEC

DRAFT

May 12, 1995

MEMORANDUM

SUBJECT: Proposed method for comparing site sample values to background values for surface and subsurface soils.

I. INORGANICS

This memorandum addresses the issue of identifying contaminated sites at the Charleston Naval Base by comparing chemical concentrations in soil samples taken at the sites with concentrations in samples taken at background locations. Data from Zone H have been used to assess the utility of various statistical approaches in attempting to determine the most appropriate means of characterizing background concentrations and comparing them with concentrations at sites. Potentially contaminated sites in Zone H include ten Solid Waste Management Units (SWMUs) and sixteen Areas of Concern (AOCs). This memo documents a five-step procedure that is being developed and implemented using data from Zone H. Discussion of the methodology is followed by application of the method to lead (Pb) data from surface soils (Level 1). The lead example is included to demonstrate the procedure; results should be considered preliminary and are not intended to support risk assessment or management decisions.

A. Develop rules for dealing with nondetect (ND) data

Following guidelines presented in various USEPA documents, one-half of the sample quantitation limit (SQL) is used for nondetect values. In practice, this means using one-half of the "U" values reported by the lab and confirmed by the validator. For the metals datasets examined so far, this approach appears reasonable. Organic compounds, to be addressed later, may require a somewhat different approach.

B. Establish background for each chemical of interest

The background dataset for Zone H consists of 96 samples labelled GDH (GDHSB001-GDHSB093, GDHSB104-GDHSB107) and 8 samples labelled SGC (SGCSB001-SGCSB008), for a total of up to 104 samples at Level 1 (surface: 0-1 foot) and 63 at Level 2 (subsurface: 3-5 feet). Level 2 samples could not be taken at many locations because of a high water table. The available data values for each chemical are assembled into datasets at each soil level.

Descriptive statistics are obtained for the original data values, including frequency distribution histograms and probability plots. Results are examined and, where appropriate (i.e., histogram positively skewed, probability plot concave upward), data are transformed into natural logarithms

(LN) of their original values to provide a closer approximation to a normal distribution. Descriptive statistics of the LN-transformed data are compared to those of the originals. All of the metals datasets examined thus far have distributions that are more nearly lognormal than normal, as illustrated by the Pb_1 example included here (Figures 1-4).

It has been suggested that lognormal data indicate the presence of contamination in the samples at the high end of the range. However, "EPA's experience with environmental concentration data...suggests that a Lognormal distribution is generally more appropriate as a default statistical model than the Normal distribution, a conclusion shared by researchers at the United States Geological Survey" (EPA, 1992, page 2). The fourteen background datasets examined so far are approximately lognormal. It is more reasonable to assume that lognormal background distributions of chemical concentrations are the norm for the Naval Base, than to assume that the datasets document a background that is contaminated in comparable fashion by seven chemicals at two different depths in the soil. Nevertheless, a few potential data outliers do appear at the high end for some compounds, and it is important that the extreme values for each parameter are not considered in the estimation process so that they do not unduly influence estimated background means and variances. Normally, outliers should be removed from a dataset only in unusual circumstances, and with specific reasons for each removal. In a lognormal distribution, even apparently extreme values may fit a straight line on a probability plot of LN-transformed data. Statistical rules of thumb for outlier removal generally are based on the variance of the sample, and include methods such as the "rule of the huge error" (Taylor, 1990, page 88), in which all values greater than four standard deviations above the mean are discarded.

Because of concerns about inadvertently including contaminated samples in the background datasets, outliers here are eliminated more readily than many standard statistical guidelines would suggest. A cutoff of "mean + 2 (standard deviation)" is applied to the LN-transformed data values for each chemical. This is the same standard used in Section D.1 below, where it is discussed. Outliers are removed on a chemical-by-chemical basis, descriptive statistics are recalculated for each chemical's dataset, and the results are used to calculate the tolerance limits described in Section D.1.

C. Develop datasets for sites

Results of analyses of soil samples at the 29 identified sites are assembled into datasets for each chemical of interest at Level 1 and Level 2, for comparison with background.

D. Compare site values to background

The comparison of site to background can best be understood within the context of statistical hypothesis testing. A hypothesis test involves the creation of two hypotheses, a "null" and an "alternative" hypothesis. "In the context of background contamination at hazardous waste sites, the null hypothesis can be expressed as 'there is no difference between contaminant concentrations in background areas and onsite,' and the alternative hypothesis can be expressed as 'concentrations are higher onsite'" (RAGS, EPA, 1989a, page 4-8). Under the assumption that there is no contamination, the likelihood of any observed difference between site and background can be calculated. If the probability of the observed difference is smaller than some predetermined level, a decision is made that since the observed site samples are not likely to be from the same population as the background samples, the site is considered contaminated for a particular chemical.

There are two possible errors that can be made in this situation. The first is that a site will be considered dirty when in fact it is clean, which is called a false positive. The probability of this error, α, is controlled by specifying the level at which the null hypothesis is considered unlikely. The other possible error, the false negative rate, β, can be seen as the probability of concluding from a test that no difference exists when in reality such a difference does exist: the site will be considered clean when in fact it is dirty. The "power" of the test ($1-\beta$), which is the complement of the false negative rate, is a measure of the strength of the conclusion that a difference does exist; it can be thought of as the probability of correctly identifying a contaminated site. The calculation of β and power is somewhat more difficult, and depends upon the magnitude of the actual differences, the size of the sample, and the form of the probability distribution for the measurement process.

Table 1: Probability of Possible Conclusions of a Hypothesis Test

	Reality	
Test	Same as Background (clean)	Greater than Background (contaminated)
Same as Background	$1-\alpha$	β
Greater than Background	α	$1-\beta$

There is a trade-off, in general, between the false positive and false negative rate, given a certain sample size. A test which rarely rejects the hypothesis of "no contamination" will be more prone to make the mistake of missing an actual difference. A test which frequently concludes that contamination is present, on the other hand, will be more likely to make the mistake of concluding that a difference arising by chance is a real difference. The total amount of error can be minimized in two ways: by increasing the sample size and by using a test which is "most powerful." The choice of the form of the hypothesis test is crucial to minimizing the total error.

EPA Region IV often suggests a "2 x background" test: If the maximum detected concentration of a chemical at a site exceeds twice the mean background level, the chemical should be considered a COPC and should be the subject of a detailed risk analysis (i.e., the chemical is a contaminant at the site). What is often not recognized is that this procedure is a statistical one, and is subject to the same errors as a hypothesis test. The problem with this approach is that background levels are never level; that is, the nature of the background data greatly affects the result of applying the "2 x background" criterion. For a normally distributed variable with a coefficient of variation (CV) of 0.25, less than 0.01% of the population is greater than twice the mean; if the CV is 1.25, 21.2% of the population exceeds the standard. In the latter case, 21.2% of the presumably uncontaminated background population would be rated contaminated by the test (false positive rate = 21.2%). Of the 14 datasets that have been examined as of the date of this memo, fully half (7) have CVs above 1.0; the range of CVs is from 0.71 to 1.41. This test neglects the information about variation which is present in the background samples, and therefore cannot be the most statistically powerful test since it does not make the most effective use of the available data.

Hypothesis tests should be suited to the type of decision that needs to be made, as well as to the type of data available. Any method for comparing site to background must be capable of detecting two different kinds of site contamination. The first type involves localized "hot spots" within the site; for example, one or two site samples out of nine or ten might test well above the highest background samples, while the rest are low or even nondetect. This situation will be modeled as a mixture of two distributions — some of the samples from a given site come from a distribution similar to the background samples while others from the same site come from a second distribution with a higher mean/median. The other type of contamination occurs when most or all of the site samples are above the mean of background samples, but none are necessarily above the high end of the background range. This situation will be modeled assuming that the distribution of site samples is similar to background, but with a higher mean/median. The first scenario will be referred to as the mixture scenario, and the second as the shift scenario. Two complementary tests are proposed for these two situations respectively — a tolerance interval test and a Wilcoxon rank sum test.

D.1. Mixture Scenario: Test Individual Samples vs. Background

Individual data values from a site can be compared to a high percentile (95th, 98th, 99th) of background values. This operation can be done parametrically by comparing to a percentile of the distribution of background values, obtained either from a probability chart of LN-transformed values or by using standard methods of estimating quantiles (e.g., Gilbert, 1987, page 175, Eqn. 13.24). It can also be done nonparametrically by comparing to a percentile of the background data values themselves, rather than to an assumed distribution of the values.

Rather than comparing site values to specific percentiles of the background data, it is possible to compare them to estimated tolerance intervals that enclose a specified percentage of the background population. A one-sided tolerance interval with 95% coverage and 95% confidence signifies that approximately 95% of individual population values fall below the upper limit, with 95% confidence. Once the interval is constructed, each site sample is compared to the upper tolerance limit (EPA, 1992, page 51). Any value that exceeds the limit is considered evidence of contamination at that point.

A roughly lognormal distribution of background values allows the use of parametric tolerance intervals, using LN-transformed values, when the nondetect percentage is low. This is the approach favored by both the Ohio Environmental Protection Agency and the Texas Natural Resource Conservation Commission to determine whether onsite contamination is greater than background. Individual sample values are compared to an upper tolerance limit that is calculated using the expression

$$\exp[X + k\,(s)]$$

where:

X = mean of LN-transformed background values
s = standard deviation of LN-transformed values
k = tolerance factor (Ohio EPA, 1991)

The tolerance factor k is obtained from tables with specified levels of α and P_0, where $(1 - P_0)$ equals the proportion of the population contained within the tolerance intervals. For a given set of α and P_0, k depends on the sample size n. For n = 63 (the sample size for Level 2 of background), k = 2.007 when $\alpha = 0.05$ and $P_0 = 0.05$ (coverage = 95%, confidence = 95%); under the same conditions of α and P_0, k = 1.917 when n = 105 (the sample size for Level 1 of background). For the sake of simplicity, a tolerance factor of k = 2 is applied to the background datasets for metals, yielding a cutoff value of

$$\text{mean} + 2\ (\text{standard deviation})$$

to determine whether a site value will be considered contaminated. In the case of a site sample contaminated with lead, for example, this method allows us to say, "We are 95% confident that this individual sample contains more lead than 95% of the population of background samples."

When a significant proportion of the samples are nondetect, it may be necessary to employ nonparametric tolerance intervals. In practice, this means using either the largest or the second largest observed background value as the standard of comparison (EPA, 1992, page 54). For a sample size of 63, using the largest background value gives coverage of over 95% with 95% confidence; for a sample size of 105, using the second largest value gives equivalent coverage and confidence levels. When nondetects reach 85-90% or more, background values may be modeled with a Poisson distribution, and Poisson tolerance limits can be constructed (EPA, 1992, page 38).

The power of this tolerance-limit test will vary based upon several factors, such as the number of samples that are assumed to have come from the distribution with the larger mean, the magnitude of the shift in the mean, and the distribution of the background samples. It also depends upon the sample size of each site and the sample size of the background. Therefore, power will depend upon the sampling strategy for each zone, and cannot be specified in a general memo. A detailed power analysis will be conducted for each zone to be included in the RFI report.

D.2. Shift Scenario: Test Entire Site vs. Background

For the situation in which the majority of samples at a site are higher than the mean background value, but none are dramatically higher, the site samples as a group must be shown to be significantly higher than the group of background samples, for contamination to be identified at the site. Figure 2.1 (enclosed) from an EPA guidance document on soils and solid media (EPA, 1989b, page 2-4) was borrowed from another document on groundwater monitoring but specifically applied to soil contamination. The upper part of the figure shows that, starting with an initial null hypothesis of "no contamination," the lower limit (confidence or tolerance) around the mean or median of the distribution of site samples must be shown to exceed the corresponding upper limit of the background distribution for the site to be considered contaminated. (As illustrated in the figure, the lower limit of the site distribution must also exceed a risk-based standard before triggering corrective action.) Depending on the nature of the data used, the upper and lower limits can be obtained using either parametric or nonparametric procedures. For a dataset with any significant number of nondetects, unfortunately, a calculated lower limit will tend to be inaccurate because it is based on the lowest data values, which must be estimated from the "U" values in the original data. Because of this limitation, the approach was rejected.

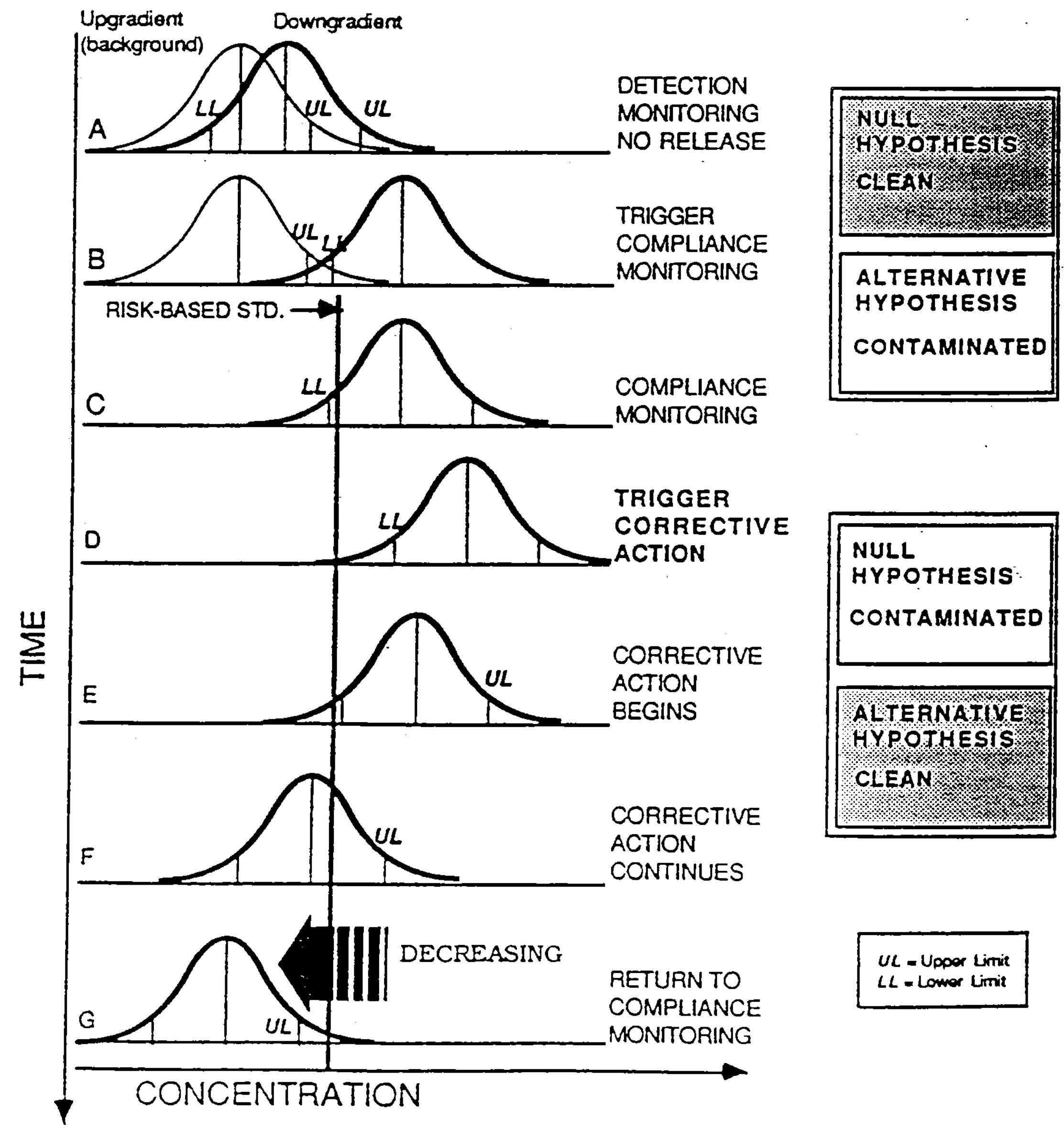
The most commonly prescribed method for comparing two populations is the *t* test, which determines whether the two population means differ significantly. The *t* test is not being used to compare site values to background because it is parametric. Although the background data values are approximately normally distributed after being LN-transformed, there is no reason to expect that the site values will be. In addition, the presence of estimated values for the nondetects calls into question the accuracy of the calculated means that are being compared.

A nonparametric counterpart to the *t* test is the Wilcoxon rank sum test, also known as the Mann-Whitney U test. Since it is nonparametric, the two datasets that are compared need not be drawn from normal or even symmetric distributions, and the test can accommodate a moderate number of nondetect values by treating them as ties (Gilbert, 1987, page 248). The method for handling nondetect values is important because it affects their ranks. "Detected but not quantified values" (J's) should receive higher ranks than nondetects. Since the ranks of the data values are evaluated and compared rather than the values themselves, the test is not sensitive to minor inaccuracies in estimated values and does not require an estimate of the mean, nor do the data values need to be LN-transformed. The Wilcoxon test is superior to some other nonparametric tests such as the sign test or the test of proportions because it takes account of differences in concentrations, and therefore has more statistical power to detect differences in those concentrations.

The Wilcoxon rank sum test operates by combining the site and background data values and ranking them by concentration. The ranks of the site samples are then compared to the background ranks. If the site ranks as a group are significantly higher than those of the background, the null hypothesis that the site and background values came from the same population is rejected at a chosen confidence level (EPA, 1992, page 46). Each group should

Methods for Evaluating the Attainment of Cleanup Standards
Volume 1: Soils and Solid Media
EPA, 1989

Figure 2.1 A Statistical Perspective of the Sequence of Ground Water Monitoring Requirements Under RCRA



(Notice that until contamination above a risk standard is documented (D) the null hypothesis is that the facility is clean. Once the facility has been proven to be in exceedance of a health criteria then the null hypothesis is that the facility is contaminated until proven otherwise (G).)

contain at least four data values. The test is available within the Minitab® statistical program for PCs.

The Wilcoxon test is very similar in power to the t-test when samples are normally distributed, and is more powerful when a large number of outlying values are to be expected.
The power of this test will vary based upon several factors, such as the magnitude of the shift in the median, and the distribution of the background samples, the sample size of each site and the sample size of the background. Therefore, power will depend upon the sampling strategy for each zone, and cannot be specified in a general memo. A detailed power analysis will be conducted for each zone to be included in the RFI report.

Summary of Section D: Choose techniques that allow the use of statistical inference. Methods must be capable of detecting situations where (a) a small number of site values are much higher than background, and (b) site values are generally higher than background. For situation (a), LN-transform all data values where appropriate to approximate normal distributions, then compare site values to an upper tolerance limit of "mean plus two standard deviations" of the background data. Where the percentage of nondetects is high, use nonparametric tolerance limits; above 85-90% nondetects, consider using Poisson tolerance limits. For situation (b), apply the Wilcoxon rank sum test to compare each group of site values to background.

E. Combine results of D.1 and D.2

Methods described in section D.1 identify individual samples with concentrations that are significantly higher than background, while the method in section D.2 identifies entire sites. If the results from either test are positive (i.e., significantly higher than background), the sample and/or site values are compared to the corresponding EPA risk-based concentration limit for soils and, where appropriate, carried forward into detailed risk assessment.

Example: Lead values at Level 1

The results of 104 analyses of background samples were assembled into a dataset and descriptive statistics were obtained for both the original and LN-transformed data, including histograms and probability plots (Figures 1-4). When the upper tolerance limit of "mean plus two standard deviations" was applied to the transformed data to identify outliers, three values were found to be above the cutoff. In terms of the original data, the tolerance limit was 143.5 ppm, while the three outliers were 172, 151, and 320 ppm. Recalculation after deletion of the outliers yielded an upper tolerance limit of 113.9 ppm, which was greater than any of the remaining data values. Figures 5-8 are histograms and probability plots of the 101 values remaining in the dataset.

Eliminating the three highest values had the following effect on parameters of the original (untransformed) data:

	Before	After
Mean	28.93	23.43
Standard deviation	40.94	22.24
CV	1.41	0.95
Skewness	4.32	1.67
Kurtosis	27.65	5.48

Parameters of the LN-transformed data changed as follows:

	Before	After
Mean	2.79	2.72
Standard deviation	1.09	1.01
CV	0.39	0.37
Skewness	-0.10	-0.40
Kurtosis	3.17	2.95

Since the greatest relative effect on the transformed data was to increase the absolute value of the skewness from 0.1 to 0.4 (away from 0.0, which is the skewness of a perfectly normal distribution), it is possible that eliminating outliers in this case was overly conservative.

Sample analysis results were assembled into datasets for individual AOCs and SWMUs, and their values were compared to the upper tolerance limit of 113.9 ppm. Eleven of the sites had values that exceeded the cutoff value:

014: 6 of 11 samples	655: 2 of 8
019: 7 of 13	666: 1 of 7
121: 9 of 11	670: 4 of 26
136: 1 of 3	684: 1 of 22
650: 4 of 9	690: 1 of 10
653: 2 of 4	

Sample values that exceed the cutoff are marked with arrows on the enclosed sample list. ("U" and "UJ" values on the list have already been divided by 2.)

Site datasets were compared to the background dataset using the Wilcoxon rank sum test (see enclosed results). At seven of the nineteen sites tested, the null hypothesis of "no significant difference" (i.e., no contamination) was rejected, indicating that overall site values were significantly higher than background. The seven sites with overall elevated values were: 014, 019, 121, 650, 653, 670, and 684. Several sites were not tested because their data values were obviously lower than background values.

The importance of using a statistical approach to comparing sites to background is evident upon examination of the results of the Wilcoxon test on AOCs 684 and 690. Although only 1 of 22 samples at AOC 684 exceeded the upper tolerance limit of the background samples, the test found that the group of site sample values was significantly higher than background at $\alpha = 0.03$. At AOC 690, 1 of 10 samples was above the upper tolerance limit, and the median data value was virtually identical to that of AOC 684 (28.15 ppm vs. 28.60 ppm); yet the test resulted in accepting the null hypothesis of "no significant difference" because the calculated difference was not significant at the prescribed level of $\alpha = 0.05$ (the test was significant at $\alpha = 0.072$). In this case, the difference in results of the two tests was probably due to the difference in the number of samples at the two sites; the larger number at AOC 684 increased the certainty of the observed differences in concentration.

The overall approach documented in this memo is considered extremely conservative for a number of reasons: (1) the number of background samples is well above the minimum recommended in various guideline documents (RAGS, EPA, 1989a, page 4-9; Ohio EPA, 1991, page 3-9), producing greater confidence in the ability to characterize background, and to distinguish background concentrations from those at sites; (2) following methodology developed in section B, high values are removed from the background datasets whether or not they are true outliers in a conventional sense, thereby lowering the total background to which the sites are compared; and (3) the use of two complementary tests increases the likelihood that any contamination will be identified and addressed further, since a positive result from either test can trigger a detailed risk assessment.

References:

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

Ohio Environmental Protection Agency, Division of Emergency and Remedial Response (1991): *How Clean is Clean.* DERR-00-RR-009

Taylor, J.K. (1990): *Statistical Techniques for Data Analysis.* Lewis Publishers, Inc., Chelsea, MI.

USEPA (1989a): *Risk Assessment Guidance for Superfund, Vol. I: Human Health Evaluation Manual (Part A).* EPA/540/1-89/002 (RAGS).

USEPA (1989b): *Methods for Evaluating the Attainment of Cleanup Standards, Vol. 1: Soils and Solid Media.* PB89-234959.

USEPA (1992): *Statistical Analysis of Ground-Water Monitoring Data at RCRA Facilities, Draft Addendum to Interim Final Guidance.*

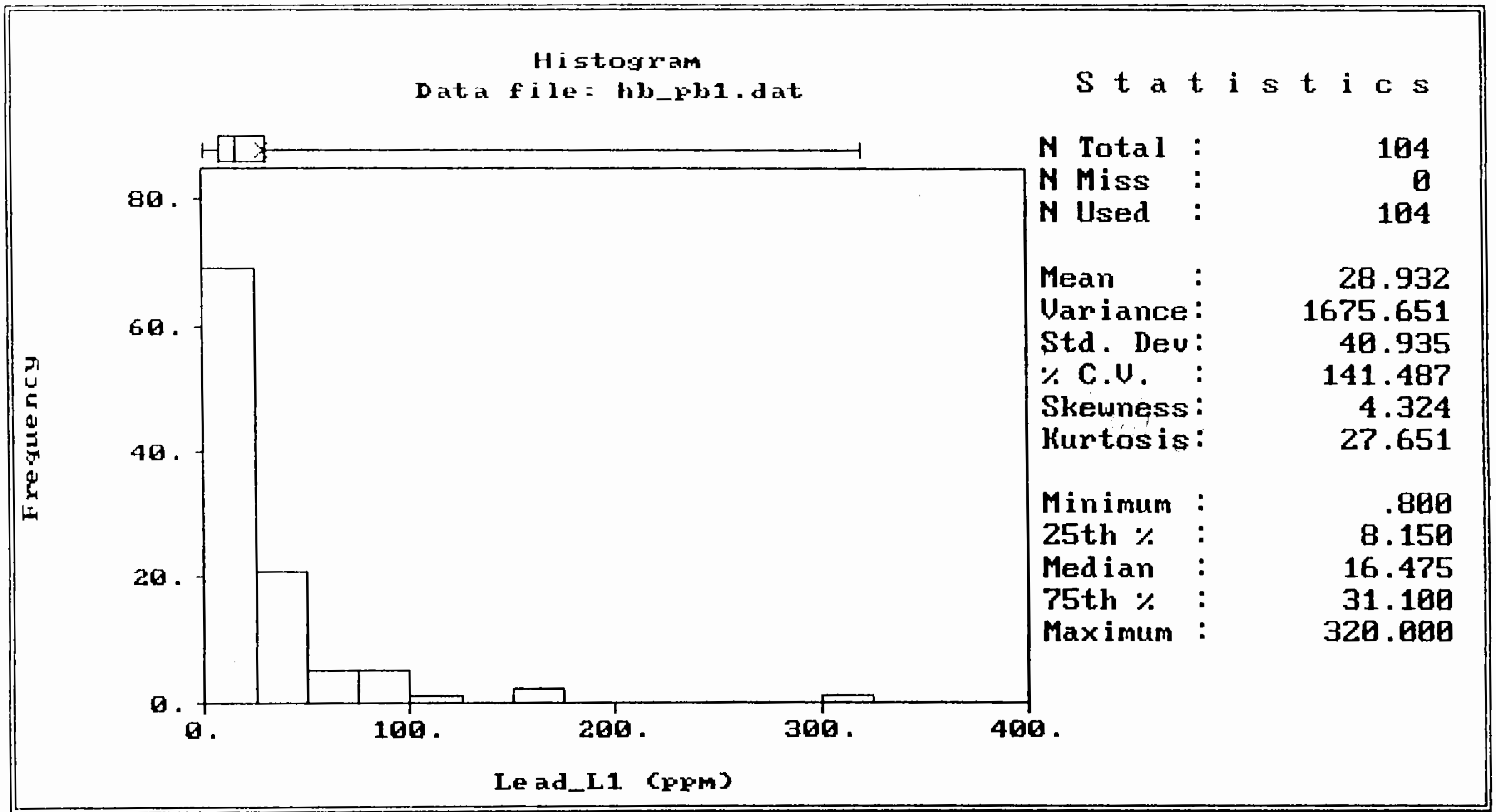


Figure 1
Frequency Histogram for Lead, Level 1
Original Values, Original Dataset

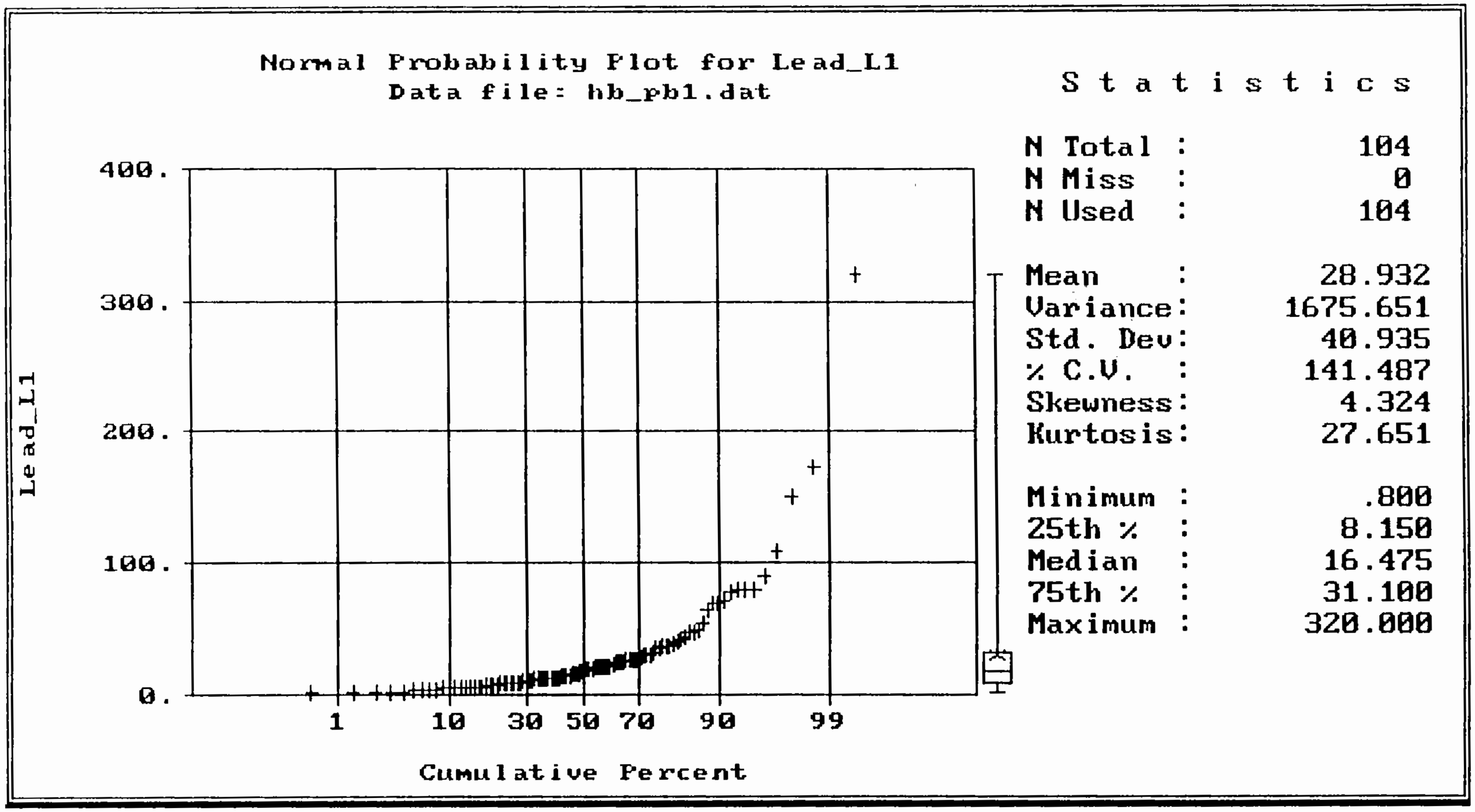


Figure 2
Probability Plot for Lead, Level 1
Original Values, Original Dataset

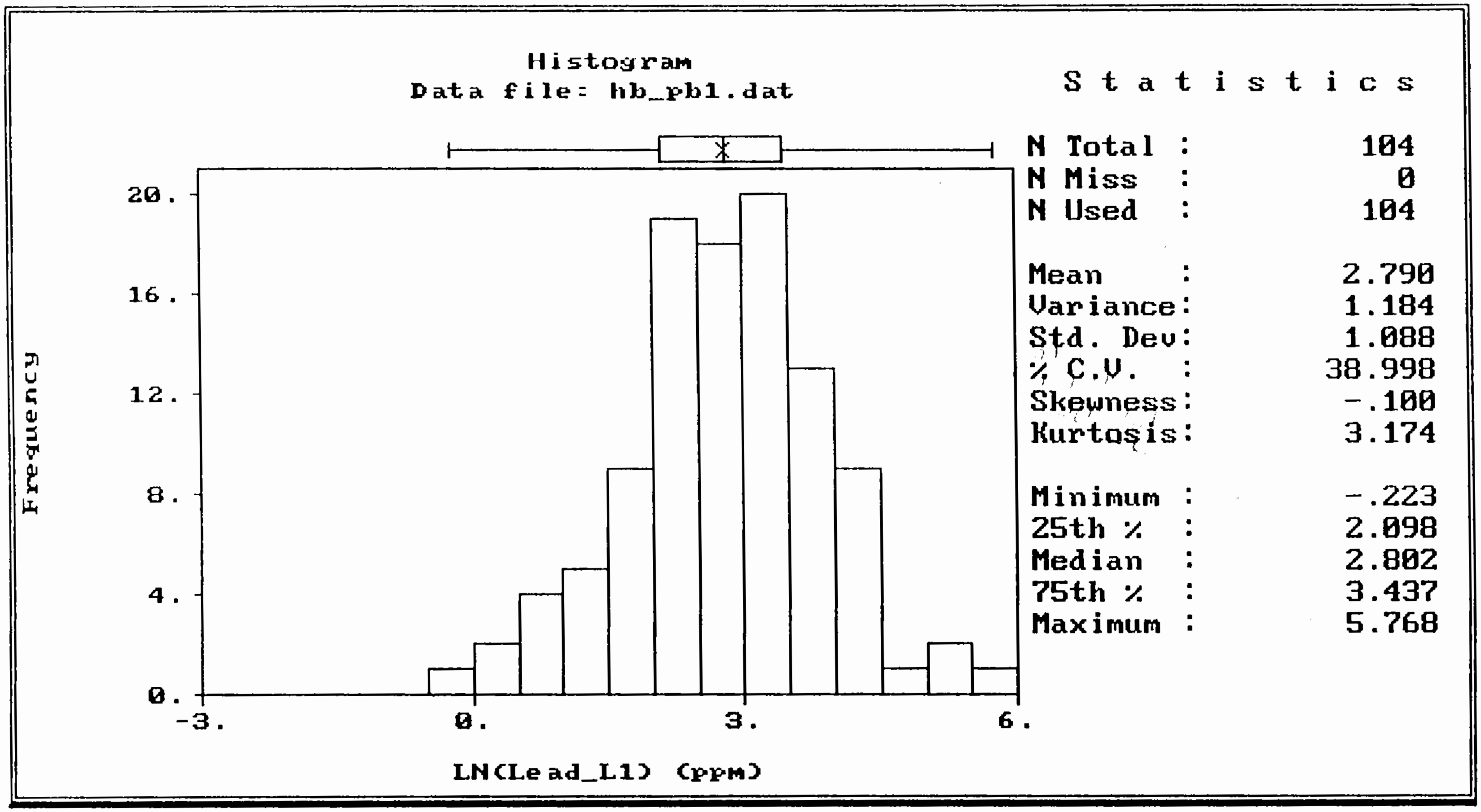


Figure 3
Frequency Histogram for Lead, Level 1
LN-Transformed Values, Original Dataset

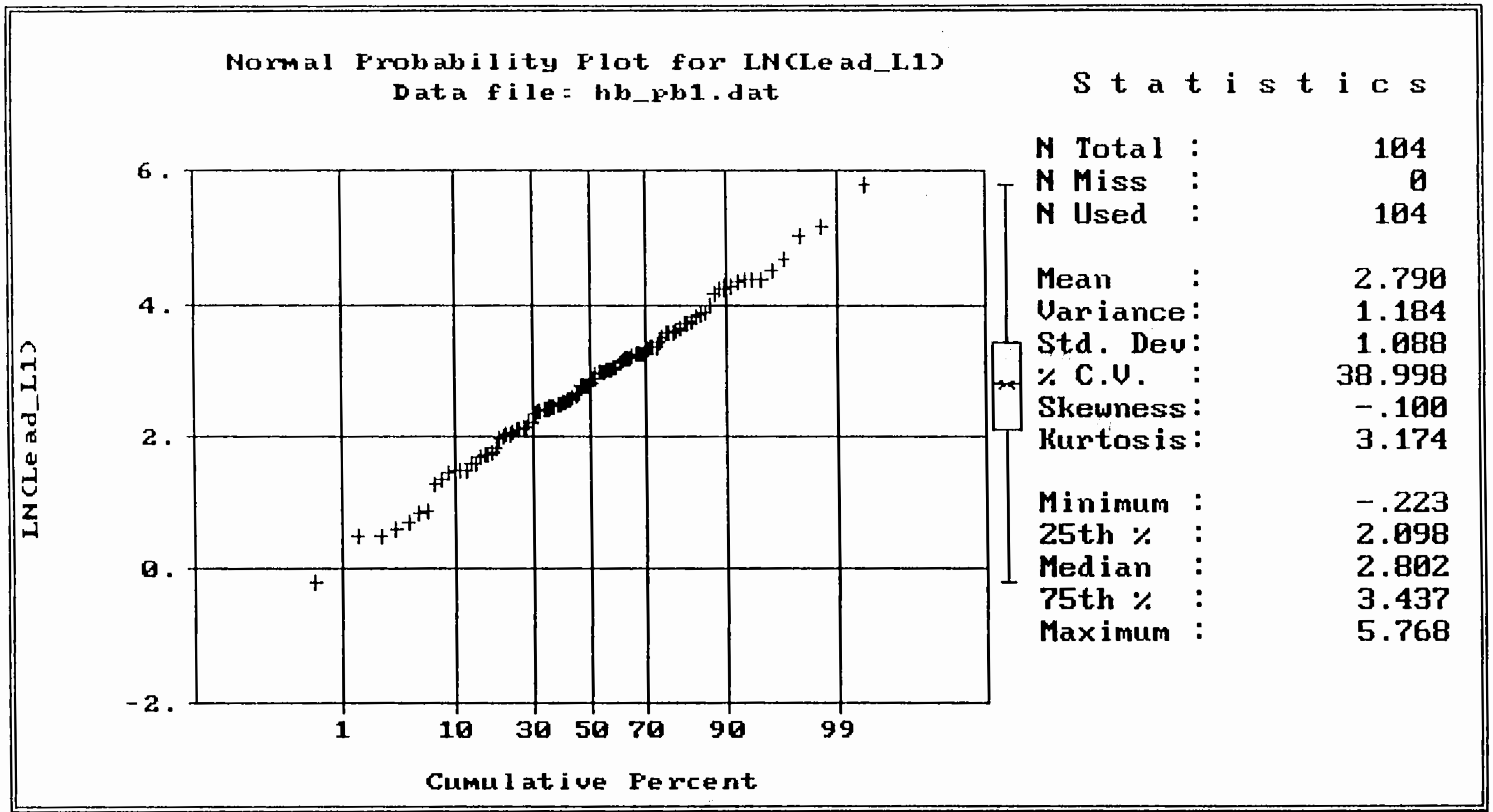


Figure 4
Probability Plot for Lead, Level 1
LN-Transformed Values, Original Dataset

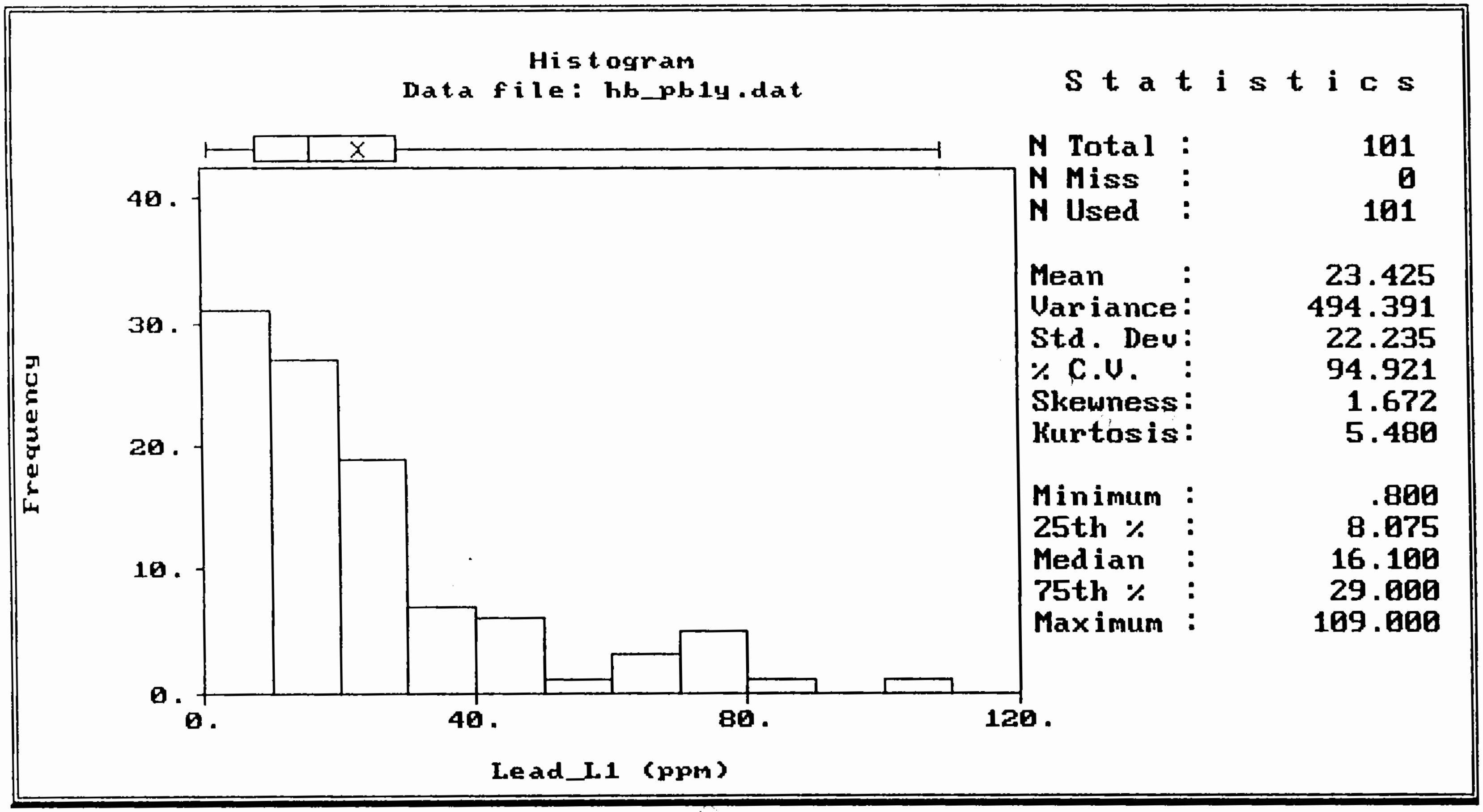


Figure 5
Frequency Histogram for Lead, Level 1
Original Values, Modified Dataset

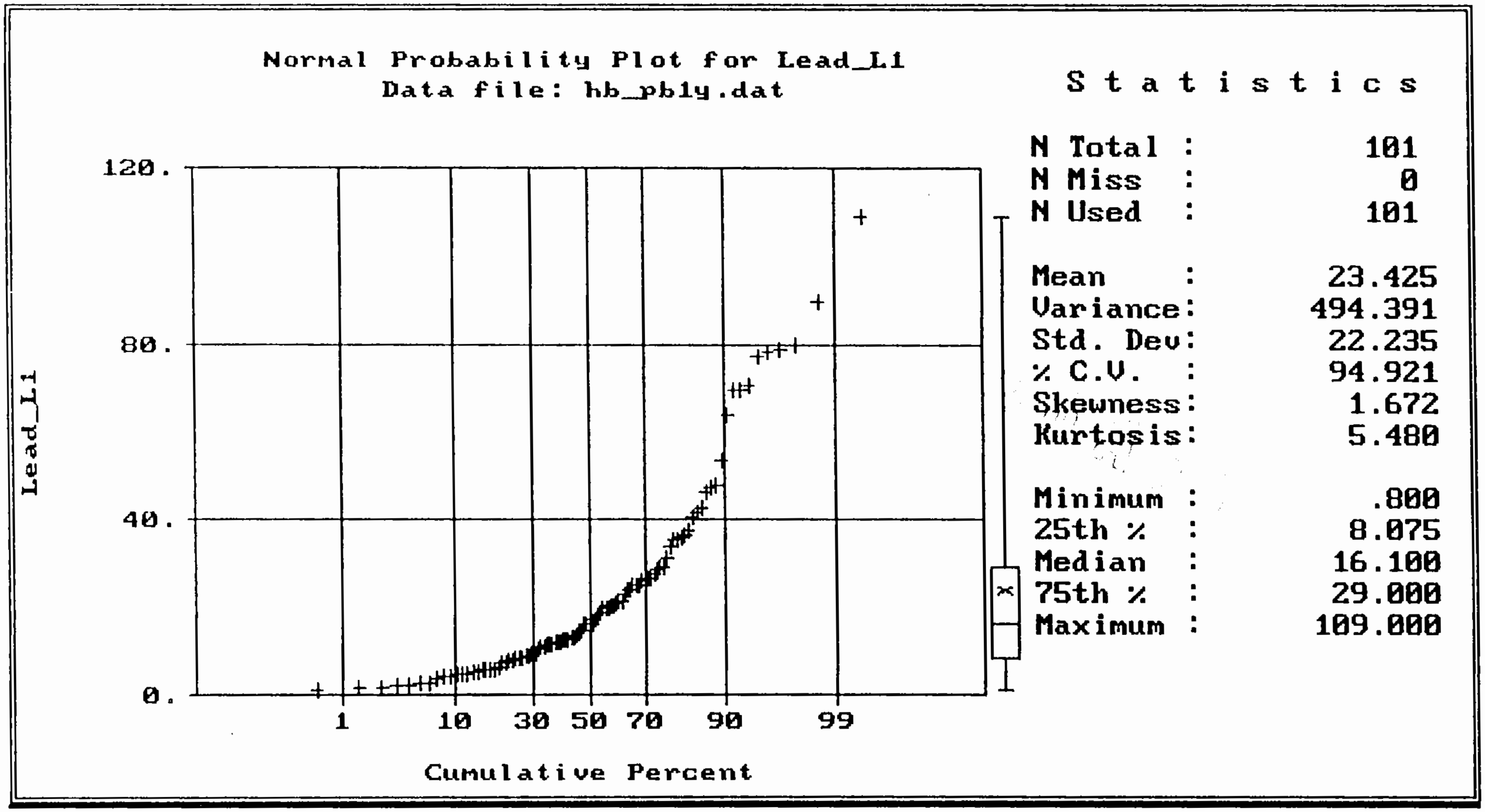


Figure 6
Probability Plot for Lead, Level 1
Original Values, Modified Dataset

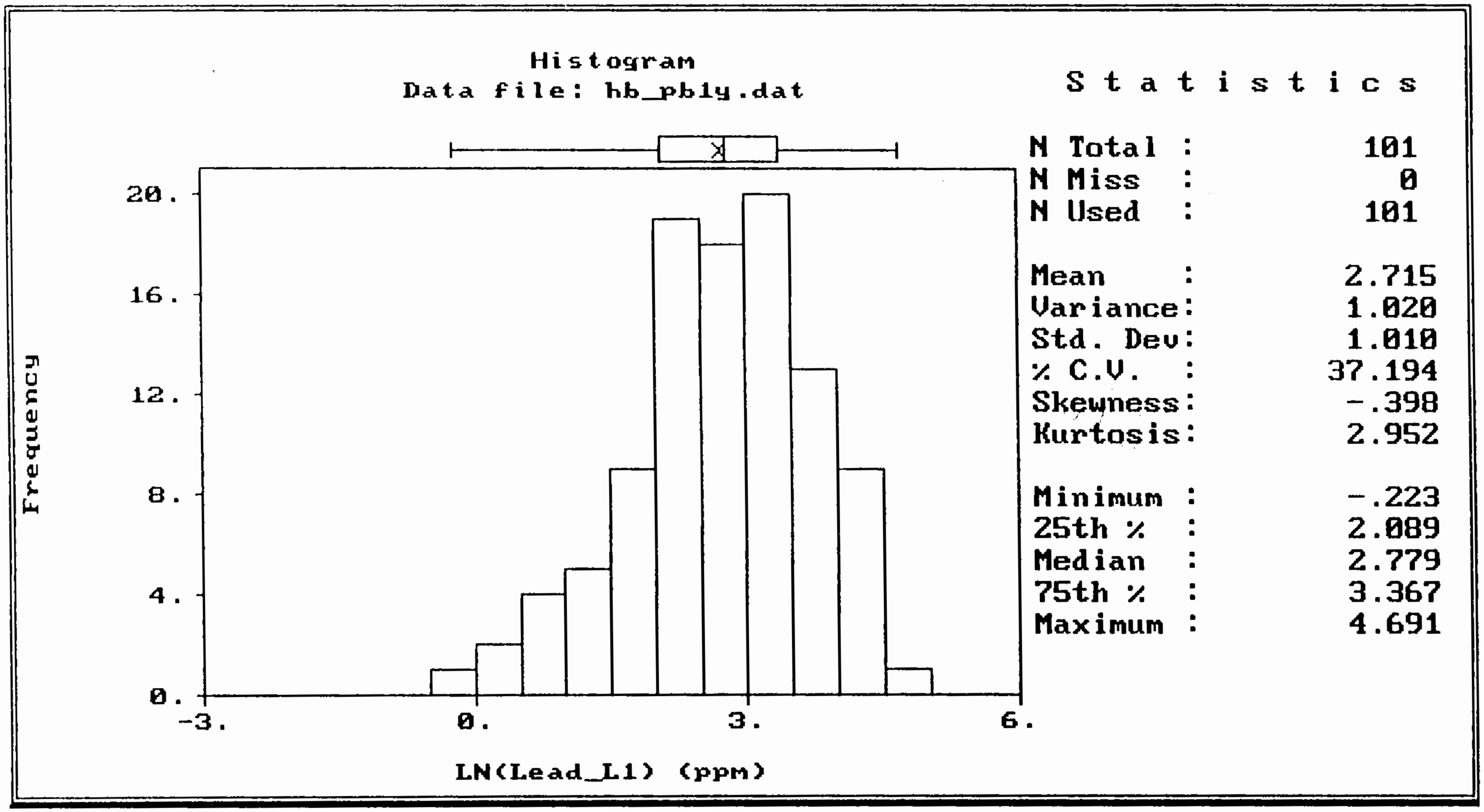


Figure 7
Frequency Histogram for Lead, Level 1
LN - Transformed Values, Modified Dataset

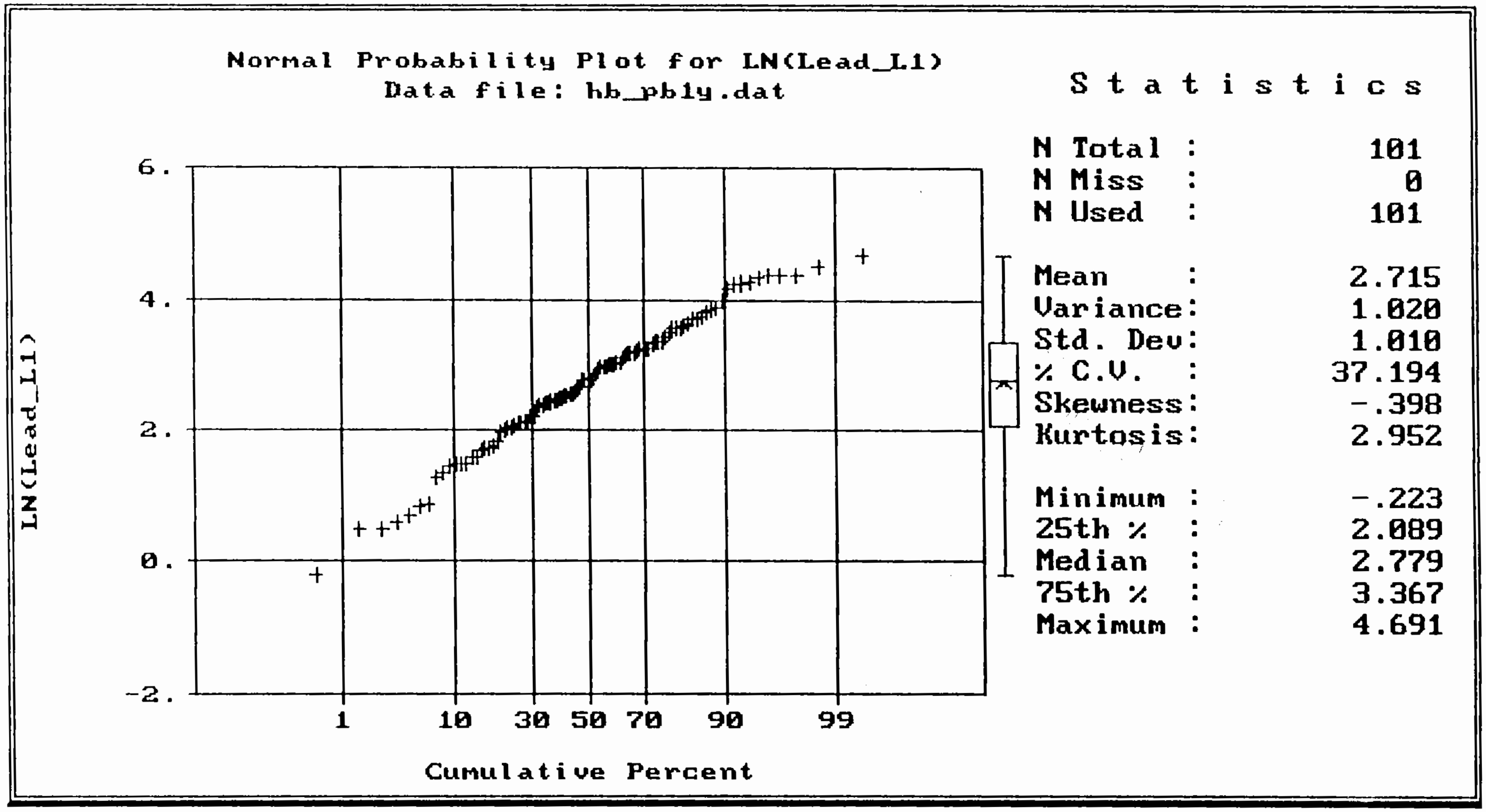


Figure 8
Probability Plot for Lead, Level 1
LN - Transformed Values, Modified Dataset

Lead in Surface Soils (Level 1)

05/03/95

Site	Matrix_id	Location	Result	Vqual
013	S	B001	8.6500	UJ
013	S	B002	3.1500	UJ
013	S	B003	9.8000	UJ
013	S	B004	54.5000	
013	S	B005	50.1000	J
013	S	B006	11.5500	U
013	S	B007	40.8000	
013	S	B008	84.7000	
013	S	B009	3.6500	U
013	S	B010	3.7000	U
013	S	B011	3.0000	U
013	S	B012	1.0000	U
013	S	B013	11.5000	
013	S	B014	2.8000	U
013	S	B015	2.6000	U
013	S	B016	15.4000	
013	S	B017	33.1000	
013	S	B018	5.9000	
013	S	B019	11.3000	
013	S	B020	14.1000	
013	S	B021	6.3000	
013	S	B022	45.5000	
013	S	B023	9.2500	U
014	S	B001	44.5000	
014	S	B002	44.6000	
014	S	B004	72.5000	
014	S	B005	→ 915.0000	
014	S	B010	→ 656.0000	
014	S	B011	→ 134.0000	
~~014~~	~~S~~	~~B011~~	~~130.0000~~	
015	S	B001	21.0000	J
015	S	B002	3.6500	U
015	S	B003	21.3000	
015	S	B004	83.7000	J
017	S	B001	19.6000	
017	S	B002	4.9000	U
017	S	B003	29.4000	
017	S	B004	30.0000	
017	S	B005	2.2000	J
017	S	B006	8.4000	
017	S	B007	4.8000	
017	S	B008	6.6000	
017	S	B009	19.6000	
017	S	B010	17.2000	J
017	S	B011	41.0000	J
017	S	B012	9.4000	
017	S	B013	7.0000	
017	S	B014	6.4000	
017	S	B015	12.5500	U
017	S	B016	6.2000	
017	S	B017	6.2500	U
017	S	B018	5.9000	
017	S	B019	26.6000	

Site	Matrix_id	Location	Result	Vqual
017	S	B020	34.2000	
017	S	B021	26.3000	
017	S	B022	36.9000	
017	S	B023	20.6000	
019	S	B001	111.0000	J
019	S	B002	→ 168.0000	
019	S	B003	25.8000	J
019	S	B004	→ 381.0000	
019	S	B005	102.0000	
019	S	B006	→ 323.0000	
019	S	B007	→ 426.0000	
019	S	B008	62.8000	
019	S	B009	→ 141.0000	
019	S	B010	52.8000	
019	S	B011	→ 607.0000	
019	S	B013	3.8000	U
019	S	B014	→ 162.0000	
121	S	B001	93.5000	
121	S	B002	→ 254.0000	
121	S	B003	→ 247.0000	
121	S	B004	→ 814.0000	
121	S	B005	→ 149.0000	
121	S	B006	→ 837.0000	
121	S	B007	→ 2,770.0000	
121	S	B008	40.6000	
121	S	B009	→ 497.0000	
121	S	B010	→ 420.0000	
~~121~~	~~S~~	~~B010~~	~~656.0000~~	
121	S	B011	→ 546.0000	
136	S	B002	→ 118.0000	
136	S	B003	8.0000	U
136	S	B004	28.4000	
138	S	B001	2.8000	J
138	S	B002	21.0000	
138	S	B003	56.7000	
178	S	B001	7.6000	
178	S	B002	9.4000	
178	S	B003	2.7000	
178	S	B004	1.3000	
178	S	B005	0.6000	
178	S	B006	5.5000	
649	S	B001	98.4000	J
649	S	B002	67.0000	J
649	S	B003	27.9000	J
649	S	B004	45.2000	J
649	S	B005	8.0000	J
649	S	B006	5.6000	
649	S	B007	3.6000	
649	S	B008	12.4000	
649	S	B009	23.8000	
649	S	B010	30.2000	
650	S	B001	→ 147.0000	J
650	S	B002	→ 121.0000	J
650	S	B003	57.0000	J
650	S	B004	→ 175.0000	J
650	S	B005	101.0000	

Site	Matrix_id	Location	Result	Vqual
650	S	B006	→ 347.0000	
650	S	B007	7.7000	
650	S	B009	73.8000	
650	S	B010	7.4000	U
653	S	B001	→ 561.0000	
653	S	B002	38.2000	J
653	S	B003	→ 638.0000	
653	S	B004	50.7000	J
654	S	B001	3.1000	U
654	S	B002	3.2500	U
654	S	B003	4.2000	U
654	S	B005	32.7000	
654	S	B006	5.2500	U
654	S	B007	5.8500	U
655	S	B001	→ 215.0000	
655	S	B002	4.4500	U
655	S	B003	7.1000	U
655	S	B004	2.3500	U
655	S	B005	→ 158.0000	
655	S	B006	5.8500	U
655	S	B007	4.5000	U
655	S	B008	23.5000	
656	S	B001	3.4000	UJ
656	S	B002	12.5500	UJ
656	S	B003	4.8000	UJ
656	S	B004	11.4500	UJ
656	S	B005	40.0000	
656	S	B006	37.4000	
656	S	B007	7.4500	UJ
656	S	B008	8.7500	UJ
656	S	B009	34.9000	
659	S	B001	2.4000	
659	S	B002	4.3000	
659	S	B003	6.9000	
659	S	B004	12.1000	
660	S	B001	2.2000	
660	S	B002	5.6000	
660	S	B003	3.9000	
660	S	B004	3.8000	
660	S	B005	6.2000	
660	S	B006	3.0000	
660	S	B007	2.1000	
660	S	B008	3.9000	
662	S	B001	4.0000	J
662	S	B002	3.9000	J
662	S	B003	4.3000	
662	S	B004	3.9000	J
663	S	B001	22.6000	
663	S	B002	37.5000	
663	S	B004	59.3000	
663	S	B005	69.3000	
663	S	B006	0.6500	U
663	S	B007	43.3000	
665	S	B001	4.3000	
665	S	B002	9.5000	
665	S	B003	11.4000	

Site	Matrix_id	Location	Result	Vqual
665	S	B004	51.4000	
666	S	B001	3.2000	J
666	S	B002	→ 118.0000	J
666	S	B003	3.6000	J
666	S	B004	20.0000	J
666	S	B005	92.2000	J
666	S	B006	4.8000	
666	S	B007	0.6000	
667	S	B001	11.9000	J
667	S	B002	11.4000	J
667	S	B003	6.8000	J
667	S	B004	3.8000	J
670	S	B001	21.0000	
670	S	B002	94.2000	
670	S	B004	4.7500	UJ
670	S	B005	35.3000	J
670	S	B006	39.6000	J
670	S	B007	5.0500	UJ
670	S	B008	20.4000	J
670	S	B009	18.2000	J
670	S	B010	26.8000	J
670	S	B011	3.8500	UJ
670	S	B012	→ 871.0000	J
670	S	B013	4.6000	U
670	S	B014	44.0000	J
670	S	B015	18.6000	J
670	S	B016	68.8000	
670	S	B017	8.1500	U
670	S	B018	45.0000	
670	S	B019	41.8000	
670	S	B020	6.1000	U
670	S	B021	7.1000	U
670	S	B022	51.2000	
~~670~~	~~S~~	~~B022~~	~~94.1000~~	
670	S	B023	→ 20,900.0000	
~~670~~	~~S~~	~~B023~~	~~411.0000~~	
670	S	B024	63.1000	
670	S	B025	→ 133.0000	
670	S	B026	→ 1,690.0000	
~~670~~	~~S~~	~~B026~~	~~81.2000~~	
670	S	B027	7.0000	U
684	S	B002	10.3000	
684	S	B003	75.9000	
684	S	B004	67.9000	
684	S	B005	61.6000	
684	S	B007	24.8000	
684	S	B008	11.2000	
~~684~~	~~S~~	~~B008~~	~~12.6000~~	
~~684~~	~~S~~	~~B008~~	~~12.2000~~	
684	S	B009	38.9000	
684	S	B010	→ 117.0000	
684	S	B011	46.4000	
684	S	B012	4.1400	
684	S	B013	8.4700	
684	S	B014	35.1000	
684	S	B015	47.2000	

Site	Matrix_id	Location	Result	Vqual
684	S	B016	3.9600	
684	S	B017	10.2000	
684	S	B018	22.3000	
684	S	B026	50.8000	
684	S	B027	32.4000	
684	S	B028	21.6000	
~~684~~	~~S~~	~~B028~~	~~16.5000~~	
684	S	B029	16.0000	
684	S	B030	5.7500	U
684	S	B031	43.0000	
690	S	B001	27.6000	
690	S	B002	28.7000	
690	S	B003	1.2500	U
690	S	B004	30.7000	
690	S	B005	94.1000	
690	S	B006	24.9000	
690	S	B007	9.0000	
690	S	B008	12.9000	
690	S	B009	33.7000	
690	S	B010	→ 118.0000	
GDH	S	B092	12.0000	
GDH	S	B001	69.5000	
~~GDH~~	~~S~~	~~B001~~	~~73.0000~~	
GDH	S	B002	47.8000	
GDH	S	B003	45.9000	
GDH	S	B004	7.8000	U
GDH	S	B005	29.2000	
GDH	S	B006	5.7000	
GDH	S	B007	74.0000	
GDH	S	B008	22.9000	
GDH	S	B009	17.3000	
GDH	S	B010	31.1000	J
~~GDH~~	~~S~~	~~B010~~	~~45.9000~~	
GDH	S	B011	41.3000	
GDH	S	B012	40.3000	
GDH	S	B013	26.2000	
GDH	S	B014	78.5000	
GDH	S	B015	→ 172.0000	
GDH	S	B016	29.2000	
GDH	S	B017	11.6000	U
GDH	S	B018	12.2000	U
GDH	S	B019	5.6500	U
GDH	S	B020	3.6000	U
GDH	S	B021	11.0000	U
GDH	S	B022	35.0000	
GDH	S	B023	26.3000	
~~GDH~~	~~S~~	~~B023~~	~~33.5000~~	
GDH	S	B024	47.0000	
GDH	S	B025	13.7000	
GDH	S	B026	53.3000	
GDH	S	B027	109.0000	
GDH	S	B028	26.3000	
GDH	S	B029	8.7000	
GDH	S	B030	12.6000	
GDH	S	B031	10.9000	
GDH	S	B032	70.5000	

Site	Matrix_id	Location	Result	Vqual
GDH	S	B033	15.8000	
GDH	S	B034	28.4000	
GDH	S	B035	37.4000	
GDH	S	B036	20.6000	
GDH	S	B037	27.5000	J
GDH	S	B038	36.2000	
GDH	S	B039	21.1000	
GDH	S	B040	4.3000	J
~~GDH~~	~~S~~	~~B040~~	~~2.6000~~	~~J~~
GDH	S	B041	33.6000	
GDH	S	B042	0.8000	U
GDH	S	B043	12.6000	
GDH	S	B044	4.4000	
GDH	S	B045	11.7000	
GDH	S	B046	14.4000	
GDH	S	B047	11.3000	
GDH	S	B048	35.3000	
GDH	S	B049	24.6000	
GDH	S	B050	8.1500	U
GDH	S	B051	7.2000	
GDH	S	B052	13.1000	
GDH	S	B053	11.4000	
GDH	S	B054	19.5000	
GDH	S	B055	16.8000	
GDH	S	B056	24.0000	
GDH	S	B057	5.8500	U
GDH	S	B058	24.6000	J
GDH	S	B059	63.8000	
GDH	S	B060	23.7000	
GDH	S	B061	9.3000	
~~GDH~~	~~S~~	~~B061~~	~~10.4000~~	
GDH	S	B062	25.0000	
GDH	S	B063	→ 151.0000	
GDH	S	B064	79.8000	
GDH	S	B065	21.2000	J
GDH	S	B066	4.5000	
GDH	S	B067	77.1000	
GDH	S	B068	19.8000	
GDH	S	B069	4.4000	
GDH	S	B071	11.5000	
GDH	S	B072	20.3000	
GDH	S	B074	4.6000	J
GDH	S	B075	89.7000	
GDH	S	B076	6.3000	
GDH	S	B077	7.4000	
GDH	S	B078	18.8000	
GDH	S	B079	69.6000	
GDH	S	B080	42.7000	
GDH	S	B081	9.2000	
GDH	S	B082	15.7000	
GDH	S	B083	7.6000	
GDH	S	B084	13.6000	U
GDH	S	B085	16.1500	U
GDH	S	B086	1.6000	
GDH	S	B087	8.7000	
GDH	S	B088	19.6000	

Site	Matrix_id	Location	Result	Vqual
GDH	S	B089	20.3000	
GDH	S	B090	41.3000	
GDH	S	B091	8.2000	
GDH	S	B093	4.9000	
GDH	S	B104	16.1000	
GDH	S	B105	17.9000	
GDH	S	B106	→ 320.0000	
~~GDH~~	~~S~~	~~B106~~	~~115.0000~~	
GDH	S	B107	2.3000	
SGC	S	B001	4.9000	U
SGC	S	B002	2.0000	J
SGC	S	B003	8.0500	U
SGC	S	B004	1.6000	J
SGC	S	B005	5.5000	
SGC	S	B006	10.5000	
SGC	S	B008	25.7000	

RESULTS OF WILCOXON RANK SUM TESTS: LEAD, LEVEL 1

```
MTB > Mann-Whitney 95.0 'Pb1_013' 'Pb1_BG';
SUBC>   Alternative 1.

Mann-Whitney Confidence Interval and Test

Pb1_013   N =  23   Median =      9.80
Pb1_BG    N = 101   Median =     16.10
Point estimate for ETA1-ETA2 is    -4.55
95.0 pct c.i. for ETA1-ETA2 is (-10.50,1.00)
W = 1180.5
Test of ETA1 = ETA2  vs.  ETA1 g.t. ETA2
Cannot reject since W is l.t. 1437.5
```

ACCEPT

```
MTB > Mann-Whitney 95.0 'Pb1_014' 'Pb1_BG';
SUBC>   Alternative 1.

Mann-Whitney Confidence Interval and Test

Pb1_014   N =  11   Median =     134.0
Pb1_BG    N = 101   Median =      16.1
Point estimate for ETA1-ETA2 is    112.9
95.1 pct c.i. for ETA1-ETA2 is (46.8,278.9)
W = 1108.0
Test of ETA1 = ETA2  vs.  ETA1 g.t. ETA2 is significant at 0.0000
The test is significant at 0.0000 (adjusted for ties)
```

REJECT

```
MTB > Mann-Whitney 95.0 'Pb1_015' 'Pb1_BG';
SUBC>   Alternative 1.

Mann-Whitney Confidence Interval and Test

Pb1_015   N =   4   Median =     21.15
Pb1_BG    N = 101   Median =     16.10
Point estimate for ETA1-ETA2 is    4.67
95.1 pct c.i. for ETA1-ETA2 is (-15.16,43.40)
W = 241.0
Test of ETA1 = ETA2  vs.  ETA1 g.t. ETA2 is significant at 0.3167
The test is significant at 0.3166 (adjusted for ties)
```

Cannot reject at alpha = 0.05 **ACCEPT**

```
MTB > Mann-Whitney 95.0 'Pb1_017' 'Pb1_BG';
SUBC>   Alternative 1.

Mann-Whitney Confidence Interval and Test

Pb1_017   N =  23    Median =      12.55
Pb1_BG    N = 101    Median =      16.10
Point estimate for ETA1-ETA2 is     -2.60
95.0 pct c.i. for ETA1-ETA2 is (-8.70,2.60)
W = 1296.5
Test of ETA1 = ETA2  vs.  ETA1 g.t. ETA2
Cannot reject since W is l.t. 1437.5
```

ACCEPT

```
MTB > Mann-Whitney 95.0 'Pb1_019' 'Pb1_BG';
SUBC>   Alternative 1.

Mann-Whitney Confidence Interval and Test

Pb1_019   N =  13    Median =      141.0
Pb1_BG    N = 101    Median =       16.1
Point estimate for ETA1-ETA2 is     114.7
95.0 pct c.i. for ETA1-ETA2 is (61.9,159.7)
W = 1258.0
Test of ETA1 = ETA2  vs.  ETA1 g.t. ETA2 is significant at 0.0000
The test is significant at 0.0000 (adjusted for ties)
```

REJECT

```
MTB > Mann-Whitney 95.0 'Pb1_121' 'Pb1_BG';
SUBC>   Alternative 1.

Mann-Whitney Confidence Interval and Test

Pb1_121   N =  10    Median =      458.5
Pb1_BG    N = 101    Median =       16.1
Point estimate for ETA1-ETA2 is     418.3
95.0 pct c.i. for ETA1-ETA2 is (235.6,541.6)
W = 1049.0
Test of ETA1 = ETA2  vs.  ETA1 g.t. ETA2 is significant at 0.0000
The test is significant at 0.0000 (adjusted for ties)
```

REJECT

```
MTB > Mann-Whitney 95.0 'Pb1_138+667' 'Pb1_BG';
SUBC>   Alternative 1.

Mann-Whitney Confidence Interval and Test

Pb1_138+667   N =   7    Median =      11.40
Pb1_BG        N = 101    Median =      16.10
Point estimate for ETA1-ETA2 is      -4.90
95.1 pct c.i. for ETA1-ETA2 is (-17.01,4.00)
W = 289.5
Test of ETA1 = ETA2  vs.  ETA1 g.t. ETA2
Cannot reject since W is l.t. 381.5
```

ACCEPT

```
MTB > Mann-Whitney 95.0 'Pb1_178' 'Pb1_BG';
SUBC>   Alternative 1.

Mann-Whitney Confidence Interval and Test

Pb1_178   N =   6    Median =     4.100
Pb1_BG    N = 101    Median =    16.100
Point estimate for ETA1-ETA2 is     -11.750
95.1 pct c.i. for ETA1-ETA2 is (-25.003,-3.803)
W = 98.0
Test of ETA1 = ETA2  vs.  ETA1 g.t. ETA2
Cannot reject since W is l.t. 324.0
```

ACCEPT

```
MTB > Mann-Whitney 95.0 'Pb1_649' 'Pb1_BG';
SUBC>   Alternative 1.

Mann-Whitney Confidence Interval and Test

Pb1_649   N =  10    Median =     25.85
Pb1_BG    N = 101    Median =     16.10
Point estimate for ETA1-ETA2 is      4.08
95.0 pct c.i. for ETA1-ETA2 is (-6.41,20.19)
W = 638.5
Test of ETA1 = ETA2  vs.  ETA1 g.t. ETA2 is significant at 0.2109
The test is significant at 0.2109 (adjusted for ties)

Cannot reject at alpha = 0.05
```

ACCEPT

```
MTB > Mann-Whitney 95.0 'Pb1_650' 'Pb1_BG';
SUBC>   Alternative 1.

Mann-Whitney Confidence Interval and Test

Pb1_650    N =   9     Median =      101.0
Pb1_BG     N = 101     Median =       16.1
Point estimate for ETA1-ETA2 is      74.8
95.0 pct c.i. for ETA1-ETA2 is (37.2,117.8)
W = 779.5
Test of ETA1 = ETA2  vs.  ETA1 g.t. ETA2 is significant at 0.0012
The test is significant at 0.0012 (adjusted for ties)
```

REJECT

```
MTB > Mann-Whitney 95.0 'Pb1_653' 'Pb1_BG';
SUBC>   Alternative 1.

Mann-Whitney Confidence Interval and Test

Pb1_653    N =   4     Median =      305.9
Pb1_BG     N = 101     Median =       16.1
Point estimate for ETA1-ETA2 is     250.9
95.1 pct c.i. for ETA1-ETA2 is (24.4,597.8)
W = 386.0
Test of ETA1 = ETA2  vs.  ETA1 g.t. ETA2 is significant at 0.0018
The test is significant at 0.0018 (adjusted for ties)
```

REJECT

```
MTB > Mann-Whitney 95.0 'Pb1_655' 'Pb1_BG';
SUBC>   Alternative 1.

Mann-Whitney Confidence Interval and Test

Pb1_655    N =   8     Median =        6.5
Pb1_BG     N = 101     Median =       16.1
Point estimate for ETA1-ETA2 is      -3.5
95.1 pct c.i. for ETA1-ETA2 is (-15.0,12.6)
W = 371.0
Test of ETA1 = ETA2  vs.  ETA1 g.t. ETA2
Cannot reject since W is l.t. 440.0
```

ACCEPT

MTB > Mann-Whitney 95.0 'Pb1_656' 'Pb1_BG';
SUBC> Alternative 1.

Mann-Whitney Confidence Interval and Test

Pb1_656 N = 9 Median = 11.45
Pb1_BG N = 101 Median = 16.10
Point estimate for ETA1-ETA2 is -2.55
95.0 pct c.i. for ETA1-ETA2 is (-13.14,6.25)
W = 439.5
Test of ETA1 = ETA2 vs. ETA1 g.t. ETA2
Cannot reject since W is l.t. 499.5

ACCEPT

MTB > Mann-Whitney 95.0 'Pb1_663' 'Pb1_BG';
SUBC> Alternative 1.

Mann-Whitney Confidence Interval and Test

Pb1_663 N = 6 Median = 40.40
Pb1_BG N = 101 Median = 16.10
Point estimate for ETA1-ETA2 is 18.88
95.1 pct c.i. for ETA1-ETA2 is (-4.25,37.64)
W = 438.0
Test of ETA1 = ETA2 vs. ETA1 g.t. ETA2 is significant at 0.0622
The test is significant at 0.0622 (adjusted for ties)

Cannot reject at alpha = 0.05

ACCEPT

MTB > Mann-Whitney 95.0 'Pb1_665' 'Pb1_BG';
SUBC> Alternative 1.

Mann-Whitney Confidence Interval and Test

Pb1_665 N = 4 Median = 10.45
Pb1_BG N = 101 Median = 16.10
Point estimate for ETA1-ETA2 is -3.10
95.1 pct c.i. for ETA1-ETA2 is (-21.39,11.09)
W = 176.0
Test of ETA1 = ETA2 vs. ETA1 g.t. ETA2
Cannot reject since W is l.t. 212.0

ACCEPT

```
MTB > Mann-Whitney 95.0 'Pb1_666' 'Pb1_BG';
SUBC>   Alternative 1.
```

Mann-Whitney Confidence Interval and Test

```
Pb1_666   N =   7    Median =      4.80
Pb1_BG    N = 101    Median =     16.10
Point estimate for ETA1-ETA2 is      -4.20
95.1 pct c.i. for ETA1-ETA2 is (-16.30,19.21)
W = 314.5
Test of ETA1 = ETA2  vs.  ETA1 g.t. ETA2
Cannot reject since W is l.t. 381.5
```

ACCEPT

```
MTB > Mann-Whitney 95.0 'Pb1_670' 'Pb1_BG';
SUBC>   Alternative 1.
```

Mann-Whitney Confidence Interval and Test

```
Pb1_670   N =  26    Median =      31.0
Pb1_BG    N = 101    Median =      16.1
Point estimate for ETA1-ETA2 is      10.7
95.0 pct c.i. for ETA1-ETA2 is (0.5,25.1)
W = 2015.5
Test of ETA1 = ETA2  vs.  ETA1 g.t. ETA2 is significant at 0.0180
The test is significant at 0.0180 (adjusted for ties)
```

REJECT

```
MTB > Mann-Whitney 95.0 'Pb1_684' 'Pb1_BG';
SUBC>   Alternative 1.
```

Mann-Whitney Confidence Interval and Test

```
Pb1_684   N =  22    Median =     28.60
Pb1_BG    N = 101    Median =     16.10
Point estimate for ETA1-ETA2 is      8.35
95.0 pct c.i. for ETA1-ETA2 is (-0.24,20.10)
W = 1649.5

Test of ETA1 = ETA2  vs.  ETA1 g.t. ETA2 is significant at 0.0300
The test is significant at 0.0300 (adjusted for ties)
```

REJECT

MTB > Mann-Whitney 95.0 'Pb1_690' 'Pb1_BG';
SUBC> Alternative 1.

Mann-Whitney Confidence Interval and Test

Pb1_690 N = 10 Median = 28.15
Pb1_BG N = 101 Median = 16.10
Point estimate for ETA1-ETA2 is 8.72
95.0 pct c.i. for ETA1-ETA2 is (-3.20,21.32)
W = 702.5
Test of ETA1 = ETA2 vs. ETA1 g.t. ETA2 is significant at 0.0718
The test is significant at 0.0718 (adjusted for ties)

Cannot reject at alpha = 0.05

ACCEPT