Cal/EPA

Department of Toxic Substances Control

P. O. Box 806
Sacramento, CA
95812-0806

301 Capitol Mall
3rd Floor
Sacramento, CA
95814

Pete Wilson
Governor

James M. Strock
Secretary for Environmental Protection

SELECTING INORGANIC CONSTITUENTS AS CHEMICALS OF POTENTIAL CONCERN AT RISK ASSESSMENTS AT HAZARDOUS WASTE SITES AND PERMITTED FACILITIES

FINAL POLICY

Prepared by:

HUMAN AND ECOLOGICAL RISK DIVISION
DEPARTMENT OF TOXIC SUBSTANCES CONTROL
CALIFORNIA ENVIRONMENTAL PROTECTION AGENCY

February 1997

Prepared by: John P. Christopher, Ph.D., D.A.B.T.
Staff Toxicologist
Human and Ecological Risk Division (HERD)
Department of Toxic Substances Control

Reviewed by: Michael J. Wade, Ph.D., D.A.B.T.
Senior Toxicologist, HERD

Brian J. Davis, Ph.D.
Staff Toxicologist, HERD

James M. Polisini, Ph.D.
Staff Toxicologist, HERD

I. Introduction

1.1 Purpose

The purpose of this policy is to provide a framework in which risk assessors may identify which inorganic constituents detected in soils at investigated sites are present at concentrations which represent contamination due to site-related activities. This is done by comparing concentrations of inorganic constituents at the site to a body of data representative of local conditions unaffected by site-related activities. For the purposes of this policy, “inorganic constituents” is limited to metals. Metals present at concentrations elevated with respect to these local conditions become chemicals of potential concern (COPC) and are carried forward into the health risk assessment. After remedial action, this same description of ambient concentrations of inorganic constituents in soil can be useful in interpreting confirmation data.

This policy is not intended to define or prescribe techniques of sampling, minimum numbers of samples, or analytical procedures. The methods described here are intended to make best use of data already available.

Following this introduction, this policy has three more parts. Section 2 presents the logical framework in which the policy is intended to be used. Section 3 gives an over-view of the two statistical methods recommended for identifying COPC. Section 4 details the steps to follow for defining the data set for ambient conditions. Appendix A describes the Wilcoxon rank sum test.

1.2 Definitions

1. “Pristine Conditions” are concentrations of metals in soils naturally occurring in locations unaffected by human activity.

2. “Ambient Conditions” are concentrations of metals in soils in the vicinity of a site but which are unaffected by site-related activities. Ambient conditions are some-times referred to as “local background”.

3. “Type I Error” is rejecting the null hypothesis when it is true. Type I error is often called a “false positive”. An example of Type I error would be identifying a metal as a COPC when its concentrations are within the range of ambient conditions.

4. “Type II Error” is accepting the null hypothesis when it is false. Type II error is often called a “false negative”. An example of Type II error would be identifying concentrations of a metal as within the range of ambient conditions, and thus not a COPC, when contamination is actually present.

2. Decision Logic

Metals eliminated as COPC are never again considered in the process of risk assessment or risk management. Thus, it is highly desirable to avoid or minimize Type II error in selection of COPC. On the other hand , if a Type I error is made, two subsequent levels of decision-making provide opportunities for correction. At the level of risk assessment, health risks due to a false positive COPC might be estimated and found to be insignificant, thus not triggering unnecessary remediation. At the level of risk management, estimated health risks due to a false positive COPC can exceed risks due to ambient conditions only slightly, a situation also unlikely to trigger unnecessary remediation. Thus, acceptable Type II error should always be less than or equal to Type I error.

3. Overview of Methods

For determining COPC, we require the use of the comparison method. To this may be added the Wilcoxon rank sum test. Both are described in general terms here. Additional details on the Wilcoxon rank sum test are give in Appendix A. When using either of the methods described here, it is necessary to follow the steps and guidance outlined below.

3.1 Comparison Method

The simplest method for identifying metals as COPC involves comparison of the highest concentration detected at the site (C_{MAX}) with a concentration representing the upper range of ambient conditions. If C_{MAX} does not exceed this value, then the metal is excluded as a COPC. If it does, the metal is carried forward into the risk assessment as a COPC. The value representing the upper range of ambient conditions may be estimated parametrically (*i.e.* making use of the underlying shape of the distribution) in most cases; or non-parametrically (no assumption about the underlying distribution).

This comparison technique has the advantage of simplicity, but it suffers from increasing Type I error (false positive) as the number of samples taken from the site increases. For example, if the 95th percentile is selected to represent the upper range of ambient concentrations, then 5% of any group of samples from a truly ambient population will exceed the 95th quantile. Since a Type I error will be made if one sample exceeds the 95th percentile, and since the probability of encountering at least one sample greater than the 95th percentile increases with the number of samples collected from the site, it follows that the probability of Type I error must increase with the number of samples from the site.

Type II error (false negative) is not formally quantifiable for the comparison method. However, it is possible to minimize the number and importance of Type II errors. Their number can be reduced by selecting a value nearer to the center of the distribution of ambient conditions as the sample size for ambient conditions grows smaller. For ex-ample, with small sample sizes a 95% upper confidence limit on the arithmetic mean or the mean itself could serve as the comparator for ambient conditions. Type I errors made at the level of selection of COPC can potentially be corrected either in the risk as-sessment or via risk management.

3.2 Wilcoxon Rank Sum Test

The Wilcoxon rank sum test (Gilbert, 1987), is described in detail in Appendix A. This test may be used as an adjunct to the comparison test for selecting COPC. The Wilcoxon rank sum test examines whether measurements from one population tend to be consistently larger (or smaller) than those from another population. Performing the Wilcoxon rank sum test involves combining the two sets of concentrations from ambient conditions and from the site, ranking these values from lowest to highest, and summing the ranks for the values from the site. This sum is designated W_{RS}. For small sample sizes ($3 \leq n < 10$ for both data sets), a value W_{RS} greater than a critical value for a given level of significance indicates an upward shift in the mean, *i.e.*, the mean concentration at the site is greater than the mean for ambient conditions. In this case, the metal is retained as a COPC. If W_{RS} is less than this critical value, then the mean concentration at the site is not greater than that of the mean for ambient conditions and the metal is eliminated as a COPC. For larger sample sizes ($n \geq 10$ for both data sets), W_{RS} is used together with data on the number of tied ranks to calculate another statistic, designated Z_{RS}. If Z_{RS} is greater than a critical value for a given level of significance, then the mean concentration at the site is greater than that of the mean for ambient conditions and the metal is identified as a COPC. If Z_{RS} is less than the critical value, then the metal is excluded as a COPC.

The Wilcoxon rank sum test is a non-parametric (distribution-free) test which has the advantage of permitting formal quantification of rates of Type I and Type II errors. Such formalization is useful in the context of USEPA methods for Data Quality Objec-tives (USEPA, 1994) and Data Quality Assessment (USEPA, 1996). However, the Wil-coxon rank sum has the disadvantage of requiring more calculations than the comparison method.

3.3 Considerations of Sample Size

Multiple measurements of a metal in either ambient or site soils will describe a distribution of concentrations for that metal. When few data are available, this distribution may be described only poorly; perhaps only the central tendency may be estimated with confidence. When large data sets are available, the extremes of distributions are more likely to be adequately characterized. Depending on the size of the ambient data set and its quality, the 95th or even the 99th percentile might be an appropriate criterion for the upper range of ambient concentrations. When sample sets for ambient conditions are large, it is often possible to use an estimate of an upper percentile of ambient concentrations as the value to be compared with C_{MAX} from the site.

4.0 Details of Selecting Ambient Data Set and Selection of COPC

The basic method for identifying metals which are COPC is to compare the highest detected concentration at the site to a value representative of the upper range of the ambient distribution. When few data are available to describe ambient conditions, both the shape of the ambient distribution and its upper extremes are uncertain and the value representative of ambient conditions should be a measure of central tendency. When ambient conditions are well described, an estimate of an upper percentile of the ambient distribution may be used. In all cases, the Wilcoxon rank sum test may be used as an adjunct to the comparison method. The steps below outline a flexible process with which project teams can define ambient conditions of metals and select metals as COPC.

4.1 Step 1: Expand the data set.

The best description of ambient conditions will be obtained from the largest data set possible. Under favorable conditions, the data set describing ambient conditions may be expanded to include samples from other studies or even possibly contaminated areas. The ambient data set can be successfully expanded under the following conditions:

4.1.1 Using Previous Studies: Data from investigations performed at the same site or nearby may be combined with the ambient data set if soil types and analytical methods are generally similar. Minor differences will be identified and can be eliminated if necessary in the analysis to follow.

4.1.2: Using data from Possibly Contaminated Areas: Samples of soil must have been analyzed for many metals. Thus, areas contaminated with one metal might display ambient concentrations for others.

4.2 Step 2: Test the distribution.

The expanded data set should tested to see if it is normally distributed. This may be done using the Shapiro-Wilks test (Gilbert, 1987) or a similar test. If the test for nor-mality fails, data should be log-transformed and tested again for log-normality. Metals present at high concentrations, such as aluminum, iron, calcium, and magnesium, tend to be normally distributed, while trace metals tend to be lognormally distributed. Distribu-tions will generally fail tests for both normality and lognormality if they contain either multiple populations or a high proportion of non-detects.

4.3 Step 3: Display summary statistics for the expanded data set.

Construct a table showing for each metal the frequency of detection, range of detected values, range of sample quantitation limits, arithmetic means and standard deviations, and coefficients of variation. Typically, data drawn from just one population will display a range of detected values of no more than 2 orders of magnitude and a coeffici-ent of variation no greater than 1. When either of these conditions is not met, one must suspect that values representative of contamination have been included in the population.

4.4 Step 4: Plot concentration vs. cumulative probability.

Sort concentration data for a metal from the lowest to the highest value. Use one-half the sample quantitation limit (SQL) for results below the detection limit (“non-detects”). Construct a plot of cumulative probability vs. concentration. It is sometimes helpful to indicate on the plot which data are non-detects. If data are lognormally distribu-ted, construct plots in base 10 to facilitate cross referencing to the descriptive statistics.

When many non-detects are present, it can be useful to assign them a dummy value at or below the lowest detected value before plotting. This can remove “noise” and aid in deciding what type of distribution is present. Figures 1 and 2 present plots of the log of arsenic concentrations in groundwater vs. cumulative probability at a site. Note that equal distances on the probability axis are equal numbers of standard deviations, not equal percentages. In Figure 1 non-detects are represented as ½SQL. The breaks in the plot indicate the aparent presence of multiple distinct populations. In Figure 2 each non-detect has been replaced by a dummy value equal to the lowest detected value. The up-per portion of the distribution in Figure 2 thus consists of detected values only and shows just one apparent population. The upper tail of the distribution of arsenic concentrations is described better in Figure 2, because scatter introduced by the use of ½SQL has been eliminated.

4.5 Step 5: Identify the population nearest the origin.

If data are drawn from just one population, the cumulative probability plot will be a straight line. If multiple, overlapping populations are present, the plot will produce a gen-tle curve instead of a straight line. Gaps or Inflection points in the plot suggest multiple populations, including possible outliers which must be eliminated. The combination of the descriptive statistics and the cumulative frequency plot forms an extremely powerful and useful tool for identifying ambient conditions.

For the purpose of identifying COPC for risk assessment, ambient condi-tions are defined as the range of concentrations associated with the population nearest the origin. This definition may be performed by inspection. The population nearest the origin is selected to minimize Type II error. This is a graphical method of eliminating outliers. Following this step, it might be useful to re-test the distribution for normality or lognormality.

4.6 Step 6: Select a value to represent the upper range of ambient conditions.

Using only the data from the population nearest the origin of the cumulative proba-bility plot, a value may be selected which represents the upper range of the distribution. This should be a value which can be supported by the available data. If sample popula-

FIGURE 1

CUMULATIVE PROBABILITY PLOT OF ARSENIC CONCENTRATIONS IN GROUNDWATER AT NAVAL WEAPONS STATION SEAL BEACH: NON-DETECTS PLOTTED AS ½SQL

:

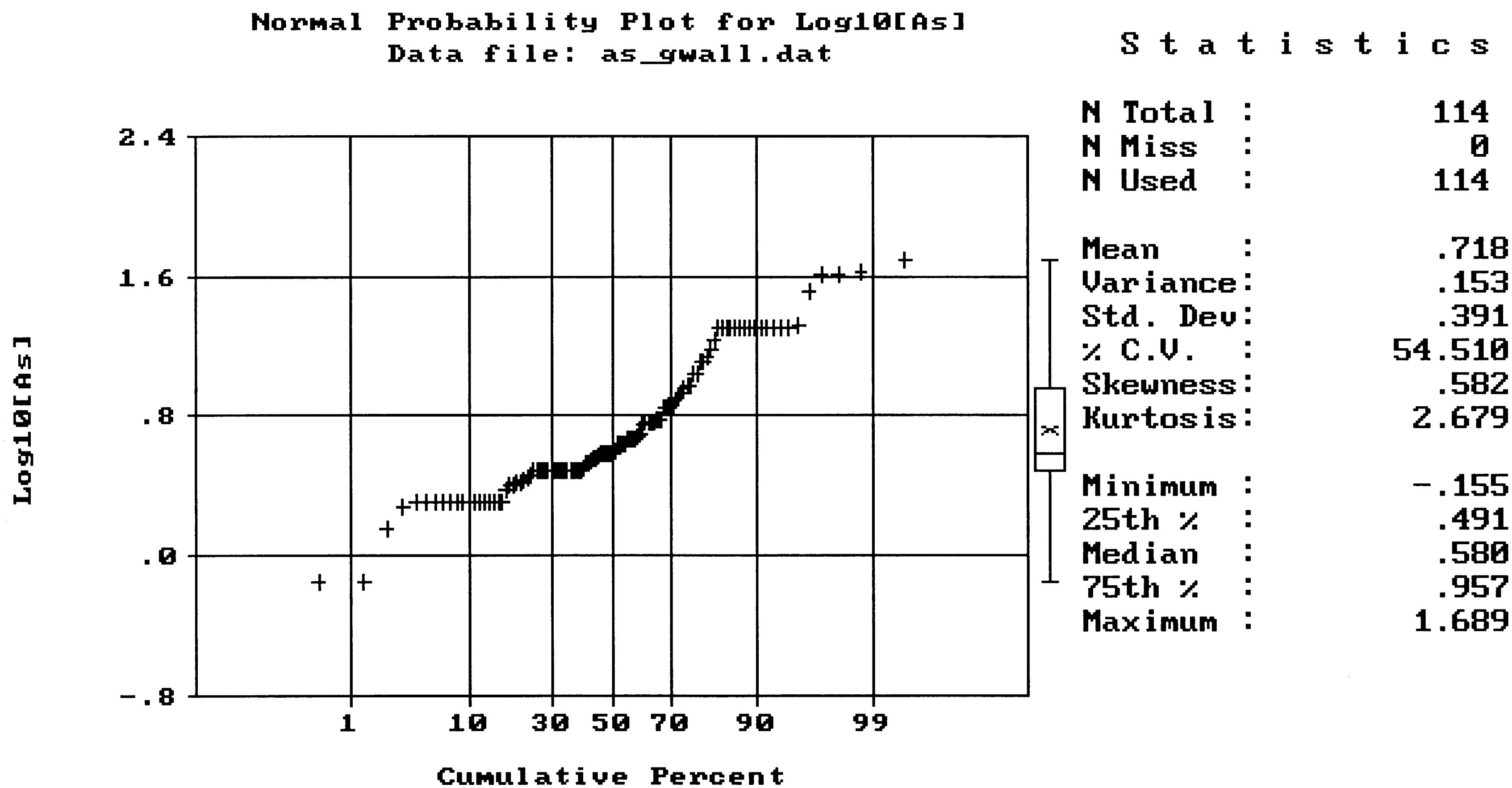
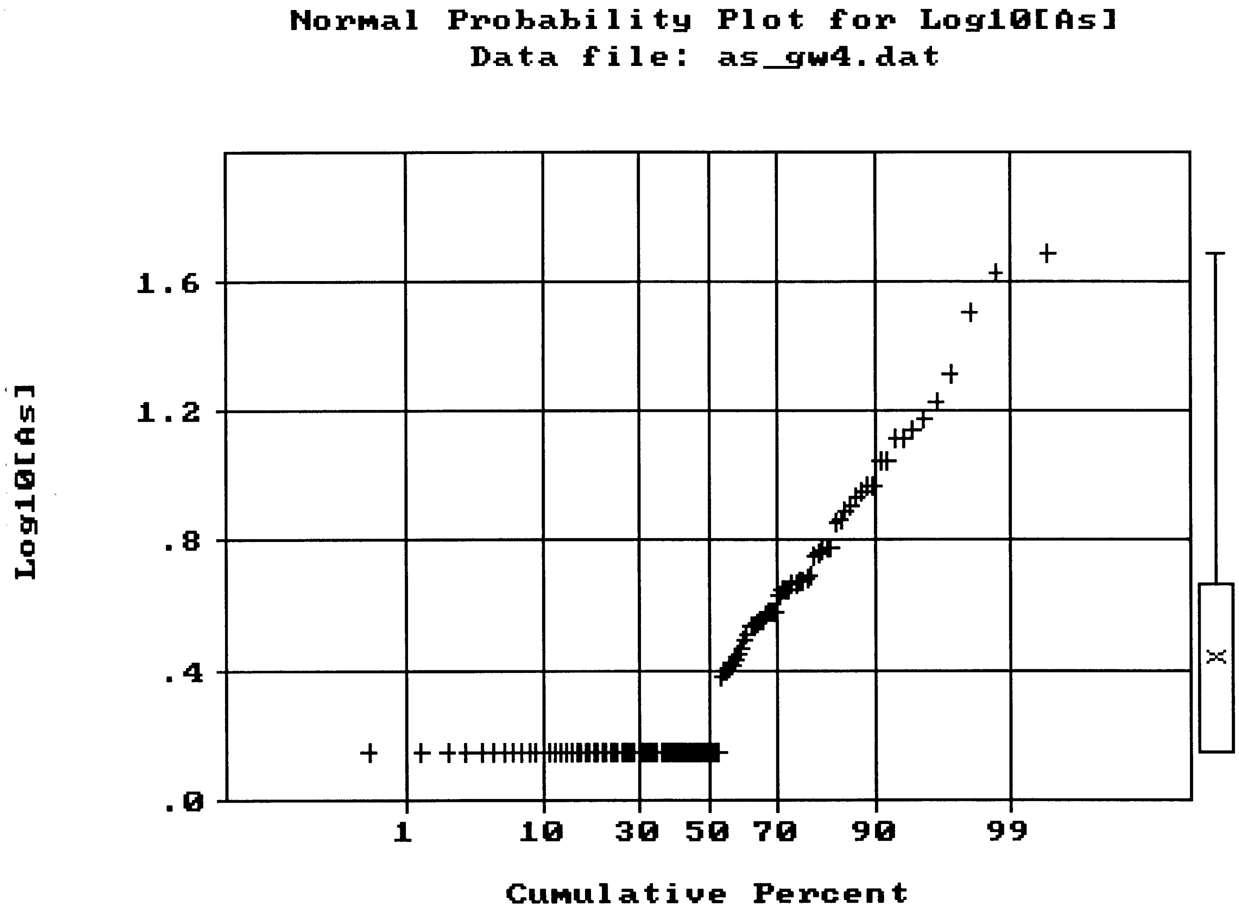


FIGURE 2

CUMULATIVE PROBABILITY PLOT OF ARSENIC CONCENTRATIONS IN GROUNDWATER AT NAVAL WEAPONS STATION SEAL BEACH: NON-DETECTS PLOTTED AS < LOWEST DETECTED VALUE



Statistics

N Total :	114
N Miss :	0
N Used :	114
Mean :	.439
Variance :	.145
Std. Dev :	.381
% C.V. :	86.845
Skewness :	1.198
Kurtosis :	3.754
Minimum :	.146
25th % :	.146
Median :	.146
75th % :	.663
Maximum :	1.689

tions are small (***n***<20), it might not be possible to estimate with confidence anything other than the central tendency, such as the arithmetic mean or an upper confidence limit about that mean. When sample sizes are larger and when the cumulative probability plot indicates that the distribution is well defined (*i.e.* little or no scatter*)*, it is acceptable to select a simple estimate of the 95th or even the 99th percentile. The selection of a representative upper quantile should be guided not by a rigid rule butr rather by the characteristics of the available data,.

Certain methods are not recommended. Upper percentiles should not be selected when data sets are small. We do not favor the uniform use of the mean plus a fixed number of standard deviations as a definition of background conditions. We do not favor the use of the upper tolerance limit or any upper confidence limit on an upper percentile as a test of back-ground, because small sample sizes inflate these estimates. We do fa-vor non-parametric statistical tests for comparing means, as long as the sample size is sufficiently large to meet the restrictions of the particular test. COPC which do not meet the restrictions for the Wilcoxon rank sum test should be retained in the risk assessment.

4.7 Step 7: Include or exclude metals as COPC.

If the highest concentration of a metal detected at a site is less than the compara-tor selected to represent the upper range of ambient conditions, then eliminate the metal as a COPC. If concentrations higher than the comparator are found, then include the metal in the risk assessment as a COPC. For those metals retained, it is often useful to examine the spatial distribution of the elevated concentrations to determine if a “hot spot” is present. If so, it could be useful to re-analyze data excluding the hot spot.

4.8 Step 8 (optional): Perform Wilcoxon rank sum test.

If many samples are collected from the site, it is possible that the Type I error rate will be unacceptable using the comparison method. In these cases, the results of the Wilcoxon rank sum test may be used as an adjunct to the comparison test for deciding whether concentrations of a metal at a site are greater than those in the ambient distribu-tion. The procedure for the Wilcoxon rank sum test is given in Appendix A.

5. References

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

U. S. Environmental Protection Agency (1994), “Guidance for Data Quality Objectives Process”, USEPA QA/G-4, September 1994.

U. S. Environmental Protection Agency (1996), “Guidance for Data Quality Assess-ment” (pre-publication copy), USEPA, QA/G-9, February 1996.

APPENDIX A

PERFORMING THE WILCOXON RANK SUM TEST

Introduction

The Wilcoxon rank sum test is presented here as described in Gilbert (1987). The test examines whether measurements from one population tend to be consistently larger (or smaller) than those from another population. The test may be performed using a hand calculator. For large data sets, computer spreadsheet software is recommended but not necessary. The test may be performed according to the steps below. An example is provided at the end of this appendix.

Assumptions and Comparison to the *t*-Test

Both the Wilcoxon rank sum test and the independent sample ***t***-test are tests of means, but the rank sum test has two main advantages. First, the two data sets need not be drawn from the same distribution. Second, the rank sum test can handle a moderate number of non-detects by treating them as ties. However, both the Wilcoxon rank sum test and the ***t***-test assume that the distributions of the two populations are identical in shape (variance), although the distributions need not be symmetric. The ***t***-test test can be modified to account for unequal variances, but no such modification exists for the rank sum test. The Gehan test, described in Gilbert (1987), is a modification of the Wilcoxon rank sum test which may also be used when non-detects are present.

Sample Size

The Wilcoxon rank sum test may be used when few samples are available for the site and the ambient data sets. The test takes slightly different forms when sample sizes are ≤ 10 or > 10.

Procedure

1. Suppose n_1 measurements represent a site and n_2 measurements represent ambient conditions. The following null hypothesis can be tested:

H_0: The populations from which n_1 and n_2 have been drawn have the same mean.

versus the following one-tailed alternative hypothesis:

H_A: The site has a higher mean than ambient conditions.

2. Select a level of significance α at which the null hypothesis may be accepted or rejected. This level is usually set at 0.05, although other levels might be selected.

3. Combine the two data sets into one with $m = n_1 + n_2$ elements. Rank these data from 1 to m in ascending order. Assign tied values a rank equal to the average of the ranks occupied by that value.

4. Sum the ranks assigned to the n_1 measurements from the site, population 1. Denote this sum by W_{RS}.

5. If either n_1 or $n_2 \leq 10$, perform a one-tailed test of H_0 versus H_A using the p-values shown in Table A-1 on page A-7. Accept H_0 and eliminate the metal as a COPC if $p > \alpha$. Accept H_A and include the metal as a COPC if $p \leq \alpha$.

6. If both n_1 and $n_2 > 10$, a normal approximation may be used. If no ties are present, compute the statistic Z_{RS} as follows:

$$Z_{RS} = \frac{W_{RS} - \dfrac{n_1(m+1)}{2}}{\left[\dfrac{n_1 n_2 (m+1)}{12}\right]^{1/2}}$$

7. If ties are present, such as NDs, compute Z_{RS} as follows:

$$Z_{RS} = \frac{W_{RS} - \dfrac{n_1(m+1)}{2}}{\left\{\dfrac{n_1 n_2}{12}\left[(m+1) - \dfrac{\sum_{j=1}^{g} t_j(t_j^2 - 1)}{m(m-1)}\right]\right\}^{1/2}}$$

where g is the number of tied groups and t_j is the number of samples with tied data in the jth group. This formulation reduces to the one shown in Step 6 in the absence of ties.

8. For a one-tailed test of H_0 versus H_A, reject H_0 and accept H_A if $Z_{RS} \geq Z_{1-a}$. Critical values of Z_{1-a} may be selected from the following table:

α	1-α	Z_{1-a}
0.10	0.90	1.282
0.05	0.95	1.645
0.025	0.975	1.960
0.01	0.99	2.327
0.001	0.999	3.080

Example Calculations

The data below are concentrations of copper in surface soil (mg Cu/kg soil), 20 values from a site and 20 from samples representative of ambient conditions:

Site:	5.9	7.4	15	18	19	19	24	31	31	34
	36	40	42	45	46	53	62	66	69	81
Ambient:	5.5	5.6	6.3	8.8	11	13	15	16	16	18
	19	20	20	22	25	30	31	50	57	73

Example 1: These data may be reformatted thus:

Copper mg/kg	Rank		Group (g)	t_j
	Site	Ambient		
5.5		1	1	1
5.6		2	2	1
5.9	3		3	1
6.3		4	4	1
7.4	5		5	1
8.8		6	6	1
11		7	7	1
13		8	8	1
15	9.5	9.5	9	2
16		11.5, 11.5	10	2
18	13.5	13.5	11	2
19	16	16, 16	12	3
20		18.5, 18.5	13	2
22		20	14	1
24	21		15	1
25		22	16	1
30		23	17	1
31	25, 25	25	18	3
34	27		19	1
36	28		20	1
40	29		21	1
42	30		22	1
45	31		23	1
46	32		24	1
50		33	25	1
53	34		26	1
57		35	27	1
62	36		28	1
66	37		29	1
69	38		30	1
73		39	31	1
81	40		32	1
W_{RS}	480		$\sum_{j=1}^{g} t_j(t_j^2 - 1)$	72

Since $n_1 > 10$, $n_2 > 10$, and some ties are present, calculate W_{RS} and Z_{RS} using Steps 4 and 7 above. Select $\alpha = 0.05$ and reject H_0 if $Z_{RS} > 1.645$. The sum of the ranks for the site W_{RS} is 496 and $m = n_1 + n_2 = 40$. Therefore, Z_{RS} may be calculated:

$$Z_{RS} = \frac{480 - \dfrac{(20)(40+1)}{2}}{\left\{\dfrac{(20)(20)}{12}\left[(40+1) - \dfrac{(4)(2)(2^2-1) + (2)(3)(3^2-1)}{(40)(40-1)}\right]\right\}^{1/2}}$$

$$Z_{RS} = 1.89$$

$Z_{RS} > 1.645$, so H_0 is rejected and H_A is accepted. It is concluded that copper is present at the site at concentrations higher than ambient conditions, so copper is retained as a chemical of potential concern for the risk assessment.

Example 2: If the data had consisted of the five lowest values from the site and the six lowest values from ambient conditions, we would have:

Site:	5.9 7.4 15 18 19
Ambient:	5.5 5.6 6.3 8.8 11 13

Copper mg/kg	Rank	
	Site	Ambient
5.5		1
5.6		2
5.9	3	
6.3		4
7.4	5	
8.8		6
11		7
13		8
15	9	
18	10	
19	11	
W_{RS}	38	

Since $n_1 \leq 10$ and $n_2 \leq 10$, calculate W_{RS} using Step 4 above. Select $\alpha = 0.05$ and reject H_0 if $p > 0.05$. The sum of the ranks for the site, W_{RS}, is 38. From Table A-1, for $n_1 = 5$ and $n_2 = 6$, $W_{RS} = 38$, $p = 0.089$. Therefore, H_0 is accepted. Copper is not present at the site at elevated concentrations with respect to ambient conditions, so it is eliminated as a COPC.

Table A-1

One-Tailed Probabilities for the Null Distribution of Wilcoxon's Rank Sum Statistic, W_{RS} [1,2]
(Entries are for $1 \le n_1 \le 4$, $3 \le n_2 \le 20$; and $5 \le n_1 \le 10$, $3 \le n_2 \le 10$.)

n_1	n_2	W_{RS}	p
1	**9**	10	0.100
	10	11	0.091
	11	12	0.083
	12	13	0.077
	13	14	0.071
	14	15	0.067
	15	16	0.062
	16	17	0.059
	17	18	0.056
	18	19	0.053
	19	19	0.100
		20	0.050
	20	20	0.095
		21	0.048
2	**3**	9	0.100
	4	11	0.067
	5	12	0.095
		13	0.048
	6	14	0.071
		15	0.036
	7	16	0.056
		17	0.028
	8	17	0.089
		18	0.044
		19	0.022
	9	19	0.073
		20	0.036
		21	0.018
	10	20	0.091
		21	0.061
		22	0.030
		23	0.015
	11	22	0.077
		23	0.051
		24	0.026
		25	0.013
2	**12**	23	0.099
		24	0.066
		25	0.044
		26	0.022
		27	0.011
	13	25	0.086
		26	0.057
		27	0.038
		28	0.019
		29	0.010
	14	26	0.100
		27	0.075
		28	0.050
		29	0.033
		30	0.017
	15	28	0.088
		29	0.066
		30	0.044
		31	0.029
		32	0.015
	16	30	0.078
		31	0.059
		32	0.039
		33	0.026
		34	0.013
	17	32	0.070
		33	0.053
		34	0.035
		35	0.023
		36	0.012
	18	33	0.084
		34	0.063
		35	0.047
		36	0.032
		37	0.021
		38	0.011
2	**19**	34	0.095
		35	0.076
		36	0.057
		37	0.043
		38	0.029
		39	0.019
		40	0.010
	20	36	0.087
		37	0.069
		38	0.052
		39	0.039
		40	0.026
		41	0.017
3	**3**	14	0.100
		15	0.05
	4	17	0.057
		18	0.029
	5	19	0.071
		20	0.036
		21	0.018
	6	21	0.083
		22	0.048
		23	0.024
		24	0.012
	7	23	0.092
		24	0.058
		25	0.033
		26	0.017
	8	25	0.097
		26	0.067
		27	0.042
		28	0.024
		29	0.012

Table A-1

One-Tailed Probabilities for the Null Distribution of Wilcoxon's Rank Sum Statistic, W_{RS} [1,2]
(Entries are for $1 \le n_1 \le 4$, $3 \le n_2 \le 20$; and $5 \le n_1 \le 10$, $3 \le n_2 \le 10$.)

n_1	n_2	W_{RS}	p
3	**9**	28	0.073
		29	0.050
		30	0.032
		31	0.018
	10	30	0.080
		31	0.056
		32	0.038
		33	0.024
		34	0.014
	11	32	0.085
		33	0.063
		34	0.044
		35	0.030
		36	0.019
		37	0.011
	12	34	0.090
		35	0.068
		36	0.051
		37	0.035
		38	0.024
		39	0.015
	13	36	0.095
		37	0.073
		38	0.055
		39	0.041
		40	0.029
		41	0.020
		42	0.012
	14	38	0.099
		39	0.078
		40	0.060
		41	0.046
		42	0.034
		43	0.024
		44	0.016
		45	0.010
3	**15**	41	0.082
		42	0.065
		43	0.050
		44	0.038
		45	0.028
		46	0.020
		47	0.013
	16	43	0.086
		44	0.069
		45	0.055
		46	0.042
		47	0.032
		48	0.024
		49	0.017
		50	0.011
	17	45	0.089
		46	0.073
		47	0.059
		48	0.046
		49	0.036
		50	0.027
		51	0.020
		52	0.014
		53	0.010
	18	47	0.092
		48	0.077
		49	0.062
		50	0.050
		51	0.040
		52	0.031
		53	0.023
		54	0.017
		55	0.012
3	**19**	49	0.095
		50	0.080
		51	0.066
		52	0.054
		53	0.044
		54	0.034
		55	0.027
		56	0.020
		57	0.015
		58	0.010
	20	51	0.098
		52	0.083
		53	0.069
		54	0.058
		55	0.047
		56	0.038
		57	0.030
		58	0.023
		59	0.018
		60	0.013
4	**4**	23	0.100
		24	0.057
		25	0.029
		26	0.014
	5	26	0.095
		27	0.056
		28	0.032
		29	0.016
	6	29	0.086
		30	0.057
		31	0.033
		32	0.019
		33	0.010

Table A-1 (continued)

One-Tailed Probabilities for the Null Distribution of Wilcoxon's Rank Sum Statistic, W_{RS} [1,2]
(Entries are for $1 \leq n_1 \leq 4$, $3 \leq n_2 \leq 20$; and $5 \leq n_1 \leq 10$, $3 \leq n_2 \leq 10$.)

n_1	n_2	W_{RS}	p
4	**7**	32	0.082
		33	0.055
		34	0.036
		35	0.021
		36	0.012
	8	35	0.077
		36	0.055
		37	0.036
		38	0.024
		39	0.014
	9	37	0.099
		38	0.074
		39	0.053
		40	0.038
		41	0.025
		42	0.017
		43	0.010
	10	40	0.094
			0.071
			0.053
			0.038
			0.027
			0.018
			0.012
	11	43	0.089
		44	0.069
		45	0.052
		46	0.039
		47	0.028
		48	0.020
		49	0.013

n_1	n_2	W_{RS}	p
4	**12**	46	0.085
		47	0.066
		48	0.052
		49	0.039
		50	0.029
		51	0.021
		52	0.015
		53	0.010
	13	49	0.082
		50	0.065
		51	0.051
		52	0.039
		53	0.030
		54	0.022
		55	0.016
		56	0.011
	14	51	0.096
		52	0.079
		53	0.063
		54	0.051
		55	0.040
		56	0.031
		57	0.023
		58	0.017
		59	0.012
	15	54	0.092
		55	0.076
		56	0.062
		57	0.050
		58	0.040
		59	0.031
		60	0.024
		61	0.018
		62	0.014
		63	0.010

n_1	n_2	W_{RS}	p
4	**16**	57	0.089
		58	0.074
		59	0.061
		60	0.050
		61	0.040
		62	0.032
		63	0.025
		64	0.019
		65	0.015
		66	0.011
	17	60	0.086
		61	0.072
		62	0.060
		63	0.049
		64	0.040
		65	0.032
		66	0.026
		67	0.020
		68	0.016
		69	0.012
	18	62	0.098
		63	0.083
		64	0.070
		65	0.059
		66	0.049
		67	0.040
		68	0.033
		69	0.027
		70	0.017
		71	0.013
		72	0.010

Table A-1 (continued)

One-Tailed Probabilities for the Null Distribution of Wilcoxon's Rank Sum Statistic, W_{RS} [1,2]
(Entries are for $1 \le n_1 \le 4$, $3 \le n_2 \le 20$; and $5 \le n_1 \le 10$, $3 \le n_2 \le 10$.)

n_1	n_2	W_{RS}	p
4	**19**	65	0.094
		66	0.081
		67	0.069
		68	0.058
		69	0.049
		70	0.041
		71	0.033
		72	0.027
		73	0.022
		74	0.018
		75	0.014
		76	0.011
	20	68	0.091
		70	0.079
		71	0.067
		72	0.057
		73	0.048
		74	0.041
		75	0.034
		76	0.028
		77	0.023
		78	0.018
		79	0.015
		80	0.011
5	**5**	35	0.075
		36	0.048
		37	0.028
		38	0.016
	6	38	0.089
		39	0.063
		40	0.041
		41	0.026
		42	0.015
5	**7**	42	0.074
		43	0.053
		44	0.037
		45	0.024
		46	0.015
	8	45	0.085
		46	0.064
		47	0.047
		48	0.033
		49	0.023
		50	0.015
	9	48	0.095
		49	0.073
		50	0.056
		51	0.041
		52	0.030
		53	0.021
		54	0.014
	10	52	0.082
		53	0.065
		54	0.050
		55	0.038
		56	0.028
		57	0.020
		58	0.014
		59	0.010
6	**6**	48	0.090
		49	0.066
		50	0.047
		51	0.032
		52	0.021
		53	0.013
6	**7**	52	0.090
		53	0.069
		54	0.051
		55	0.037
		56	0.026
		57	0.017
		58	0.011
	8	56	0.091
		57	0.071
		58	0.054
		59	0.041
		60	0.030
		61	0.021
		62	0.015
		63	0.010
	9	60	0.091
		61	0.072
		62	0.057
		63	0.044
		64	0.033
		65	0.025
		66	0.018
		67	0.013
	10	64	0.090
		65	0.074
		66	0.059
		67	0.047
		68	0.036
		69	0.028
		70	0.021
		71	0.016
		72	0.011

Table A-1 (continued)

One-Tailed Probabilities for the Null Distribution of Wilcoxon's Rank Sum Statistic, W_{RS} [1,2]
(Entries are for $1 \le n_1 \le 4$, $3 \le n_2 \le 20$; and $5 \le n_1 \le 10$, $3 \le n_2 \le 10$.)

n_1	n_2	W_{RS}	p
7	**7**	64	0.082
		65	0.064
		66	0.049
		67	0.036
		68	0.027
		69	0.019
		70	0.013
	8	68	0.095
		69	0.076
		70	0.060
		71	0.047
		72	0.036
		73	0.027
		74	0.020
		75	0.014
		76	0.010
	9	73	0.087
			0.071
			0.057
			0.045
			0.036
			0.027
			0.021
			0.016
			0.011
	10	77	0.097
			0.081
			0.067
			0.054
			0.044
			0.035
			0.028
			0.022
			0.017
			0.012
8	**8**	81	0.097
		82	0.080
		83	0.065
		84	0.052
		85	0.041
		86	0.032
		87	0.025
		88	0.019
		89	0.014
		90	0.010
	9	87	0.084
		88	0.069
		89	0.057
		90	0.046
		91	0.037
		92	0.030
		93	0.023
		94	0.018
		95	0.014
		96	0.010
	10	92	0.086
		93	0.073
		94	0.061
		95	0.051
		96	0.042
		97	0.034
		98	0.027
		99	0.022
		100	0.017
		101	0.013
		102	0.010
9	**9**	101	0.095
		102	0.081
		103	0.068
		104	0.057
		105	0.047
		106	0.039
		107	0.031
		108	0.025
		109	0.020
		110	0.016
		111	0.012
	10	107	0.091
		108	0.078
		109	0.067
		110	0.056
		111	0.047
		112	0.039
		113	0.033
		114	0.027
		115	0.022
		116	0.017
		117	0.014
		118	0.011
10	**10**	123	0.095
			0.083
			0.072
			0.062
			0.053
			0.045
			0.038
			0.032
			0.026
			0.022
			0.018
			0.014
			0.012

1. From Hollander, M., and Wolfe, D. A., *Nonparametric Statistical Methods*, Table A.5, pp. 272-282, John Wiley & Sons, New York, 1973.

2. Entries are exact values for $0.010 \leq p \leq 0.100$. For a given n_1 and n_2, entries are omitted when no higher value for W_{RS} exists, when $p > 0.100$ for all lower W_{RS}, or when $p < 0.010$ for all higher W_{RS}.