APPENDIX A

Recommended Statistical Tests

APPENDIX A
RECOMMENDED STATISTICAL TESTS

TABLE OF CONTENTS

LIST OF TABLES

APPENDIX A
RECOMMENDED STATISTICAL TESTS

Three statistical procedures are documented below. The tests are 1) the skewness test for normality, 2) the Mann-Kendall test for trend, and 3) calculation of prediction intervals. The documentation provided here is sufficient to perform the tests; however, the reader is cautioned against applying statistical procedures without thorough understanding of the theory, hypotheses and limitations of the procedures. The references cited can provide the necessary background information.

Many computer based statistical analysis programs include the skewness test (or calculation of the skewness coefficient) and the Mann-Kendall test for trend (or calculation of Kendall's Tau statistic).

A1.0 The Skewness Test

References:

George W. Snedecor and William G. Cochran. *Statistical Methods.*The Iowa State University Press. Ames, Iowa. 1980.

Harris et al. Statistical Methods for Characterizing Ground-Water Quality. *Groundwater*, Vol.25, No.2. March-April 1987.

The skewness test may be used to determine whether or not a set of independent data points are drawn from a normal distribution. The test is very simple to apply. The skewness coefficient is calculated and the value compared to a critical value found in Table A-1.

The null and alternative hypotheses for the skewness test may be stated as:

H_0 : The data may be normally distributed.

H_1 : The data are not drawn from a normal distribution.

The skewness coefficient is:

$$g = \frac{\sqrt{n} \sum_{i=1}^{n} (x_i - \overline{x})^3}{\sum_{i=1}^{n} (x_i - \overline{x})^{\frac{3}{2}}}$$

where g = skewness coefficient

n = sample size

x_i = concentration at time i

$\overline{X}$ = mean concentration

To apply the skewness test, consult Table A-1. Find the critical skew associated with sample size n. If the calculated skewness coefficient is **less** than the critical skew (from the table), then the null hypothesis **is not rejected** at the 5 or 1 percent significance level. Conclude that the data may be drawn from a normal distribution. If the calculated skewness coefficient is **greater** than the critical skew in the table, then the null hypothesis **is rejected** at the 5 or 1 percent significance level. Conclude that the data are probably not drawn from a normal distribution.

Note that Table A-1 may be used to test either positive skew, negative skew or both. To test negative skew, the critical skew is just the negative of the tabulated values at any sample size. To perform a two-tailed test, compare the calculated skewness coefficient to both the positive and negative critical skew value. The test is then performed at the 10 or 2 percent significance level.

TABLE A-1
CRITICAL SKEWNESS COEFFICIENTS

Sample Size	Percentage Points 5%	Percentage Points 1%
9	0.953	1.420
10	0.950	1.395
11	0.927	1.358
12	0.915	1.331
13	0.886	1.306
14	0.861	1.291
15	0.854	1.280
16	0.833	1.246
17	0.817	1.220
18	0.798	1.197
19	0.769	1.161
20	0.777	1.146
21	0.753	1.116
22	0.742	1.099
23	0.732	1.087
24	0.710	1.074
25	0.712 (0.711)	1.060 (1.061)
26	0.689	1.013
27	0.689	1.016
28	0.674	1.006
29	0.669	0.992
30	0.651 (0.662)	0.972 (0.986)
35	(0.621)	(0.921)
40	(0.587)	(0.869)
45	(0.558)	(0.825)
50	(0.533)	(0.787)

Adapted from Table 4 (Harris et al, 1987) . Data in parentheses from Table A-20 (Snedecor and Cochran 1980)

A2.0 Mann-Kendall Test

The Mann-Kendall test for trend is also known as Kendall's Tau test. The one sided null hypothesis is:

H_O = the X_i exhibit no trend

The one sided alternative hypotheses are:

H_+ = the X_i exhibit an upward trend

H_- = the X_i exhibit a downward trend

and the two-sided alternative is:

H = there is either an upward or a downward trend.

A good reference for Kendall's Tau statistic and the associated statistical hypothesis test is

Gibbons, Jean Dickenson. *Nonparametric Methods for Quantitative Analysis (Second Edition)* . American Sciences Press, Inc.: Columbus, Ohio. 1985.

The test for trend assumes that the data points are independent of each other. This would not be the case for monthly data, for example, if there is a seasonal pattern by months. If the seasonal variation can be removed, these procedures are applicable to the adjusted data. Kendall's Tau statistic is the nonparametric analog of the parametric test based on the regression coefficient (which assumes the normal distribution) . The asymptotic efficiency of Kendall's tau to the regression test is about 0.98 for normal distributions.

To perform the test, consider each sampling event as a pair of observations (X,Y) where Y is the sample value and X is the sampling date. List the data in chronological order and assign ranks to X and to Y independently. Rank (1) is associated with the first sampling date, X_1, and the lowest sample value (if testing for positive trend).

If the X and the Y characteristics are in perfect agreement (positive trend), the Y data should be in natural order (the X data already are). If there is perfect disagreement (negative trend), the corresponding Y data is in reverse of natural order. The Kendall Tau coefficient is a relative measure of the discrepancy between the actual observed order of the Y's and the two orders that would result from perfect association. The procedure is most easily explained by an example (drawn from Gibbons 1985):

EXAMPLE:

Suppose that n=5 and two sets of ranks are paired as follows.

X rank: 1 2 3 4 5 (monitoring dates)

Y rank: 2 3 1 4 5 (sample values)

Note that the X set is in natural order. In the resulting arrangement of Y ranks, we consider all of the possible pairs of Y ranks and score a 1 for each pair of ranks that appear in natural order and -1 for those in reverse order. We take the pairs in a systematic way, as the 2 paired with each successive rank appearing to its right, then the 3 paired with each to its right, and so on. The first pair of Y ranks, 2 followed by 3, is in natural order, so its score is 1. The second pair, 2 followed by 1, is in reverse order, so -1 is scored. The resulting scores for all possible pairs are shown in Table A-2.

Note that there are $\binom{5}{2} = 10$ possible pairs. The ratio of the total plus score, in this case 8, to the maximum,10, provides a measure of relative agreement, that is 8/10. Similarly the ratio of the total minus score to the maximum, 2/10 in this case, measures the relative disagreement. The net relative score of association is then 8/10-2/10 = 6/10, and this is the value of the Kendall Tau statistic. If we let U= the number of pairs of Y values (or ranks) in natural order (that is, the number of plus scores) and let V equal the number of Y pairs in reverse order, and let S be the difference between U and V, S = U-V , then the Tau coefficient is calculated as:

$$T = \frac{2S}{n\,(n-1)}$$

where n is the number of (X,Y) pairs (sample size). For our example, T is equal to 3/5 (0.60). The test statistic presented here is not exact if ties exist in the sample data. No ties should exist in the X series, sampling dates. Any duplicate data should be averaged. If multiple dates have the same sample value the denominator of the test statistic must be adjusted (i.e. ties exist). The adjusted test statistic is more difficult to calculate and computer analysis is recommended. Gibbons (1985) presents this test statistic in detail and how it may be calculated by hand.

TABLE A-2
Calculation of Kendall Tau Statistic

Y pair	Score	Summary Totals
2,3	1	8 plus
2,1	-1	2 minus
2,4	1	
2,5	1	
3,1	-1	
3,4	1	
3,5	1	
1,4	1	
1,5	1	
4,5	1	

The statistic T above, and the T statistic adjusted for ties, is found in most statistical software packages. Interpretation of the Tau statistic is simple. If perfect positive correlation exists T is equal to 1. If perfect negative correlation exists T is equal to -1. If no correlation exists T is equal to 0. To test the null hypothesis that no trend exists, the value of T is compared to a critical value of T found on Table A-3.

For a one-sided test for positive trend, compare the calculated value of T to the associated T in Table A-3. If the calculated value is **greater than or equal to** the table value, reject the null hypothesis of no trend (conclude that positive trend exists). For a one-sided test for negative trend, compare the calculated value of T to the negative of the associated T in Table A-3. If the calculated value is **less than or equal to** the table value, reject the null hypothesis of no trend (conclude that negative trend exists). For a two-sided test for positive or negative trend, compare the calculated value of T to both the positive and negative of the associated table value. If the calculated value is greater than +T or less than -T conclude that trend exists.

TABLE A-3
Critical values for Kendall Tau Statistic

Sample Size (n)	T (tau)
5	0.80
6	0.733
7	0.619
8	0.571
9	0.500
10	0.467
11	0.418
12	0.394
13	0.359
14	0.363
15	0.333
16	0.317
17	0.309
18	0.294
19	0.287
20	0.274
21	0.267
22	0.264
23	0.257
24	0.246
25	0.24
26	0.237
27	0.231
28	0.228
29	0.222
30	0.218

Note: The T values for $n>10$ are the right tail (or left tail) critical values for a one-sided test performed at a significance level of 0.05. For $5<n<10$, the values are the lowest T for which a one sided test, performed at a significance level of 0.05, would reject the null hypothesis. In this case, the probability associated with the T value is not exact, but is always less than 0.05. For n>30, critical values may be found from a normal probability table (See Gibbons 1985). However, T = 0.218 will always be a conservative estimate.

A3.0 PREDICTION INTERVALS

References:

U.S. Environmental Protection Agency. *Statistical Analysis of Groundwater at RCRA Facilities.* Office of Solid Waste, Waste Management Division. October, 1988 (Available from NTIS Reference Number PB 89-151-047).

Gibbons, Robert D. "Statistical Prediction Intervals for the Evaluation of Ground-Water Quality." *Ground Water.* Vol. 25. pp.455-465. 1987.

A prediction interval is a statistical interval designed to define a background concentration interval within which future measurements from the same population are likely to fall. The prediction interval can answer the question "What is the concentration associated with an allowable exceedance probability given the natural variability in the data and the sample size?". The allowable exceedance probability is recommended as 0.05 by EPA.

The prediction interval is recommended to be developed on a well by well basis. Data from multiple wells should not be aggregated.

To calculate a prediction interval the mean, $\bar{x}$, and the standard deviation s, must be calculated for the data used to form the prediction interval. Then the interval is given by

$$\bar{x} + s\sqrt{\frac{1}{m}+\frac{1}{n}}\ t_{(n-1,K,0.95)}$$

where m is the number of measurements per sampling period (i.e. 2 if duplicate data are available), and n is the number of observations in the background data, and $t_{(n-1,K,\ 0.95)}$ is found from Table A-4. The table is entered with K as the number of future observations (usually 1 if comparison is done each quarter, or 4 if comparison is done annually), and degrees of freedom, ν = n-1. If K is greater than 5 (unlikely), use the column for K = 5.

To compare new data to the prediction interval, calculate the mean of duplicate measurements or just compare the new data point to see whether it falls within the interval. If the new data is not within the prediction interval, this is statistically significant evidence of contamination.

Note that for a single future observation (i.e. one observation per quarter with quarterly comparisons), the t value may be obtained straight from the t-distribution which is tabulated in most statistical texts. Also, note that the prediction intervals are one-sided, giving a value that should not be exceeded by the future observations. If a two sided interval is required, the same procedure may be used, however Table A-4 will provide interval at the 2 α percent significance level (where α is usually 0.05).

TABLE A-4
95th Percentiles of the Bonferroni t-statistics, $t_{(\nu,\ \alpha/k)}$
(adapted from EPA, October 1988)

	k	1	2	3	4	5
ν	α/k	0.05	0.025	0.0167	0.0125	0.01
4		2.13	2.78	3.20	3.51	3.75
5		2.02	2.57	2.90	3.17	3.37
6		1.94	2.45	2.74	2.97	3.14
7		1.90	2.37	2.63	2.83	3.00
8		1.86	2.31	2.55	2.74	2.90
9		1.83	2.26	2.50	2.67	2.82
10		1.01	2.23	2.45	2.61	2.76
15		1.75	2.13	2.32	2.47	2.60
20		1.73	2.09	2.27	2.40	2.53
30		1.70	2.04	2.21	2.34	2.46
>30		1.65	1.96	2.13	2.24	2.33

ν = degrees of freedom associated with the mean square error.
k = number of comparisons
α = 0.05, the experimentwise error level