GROUNDWATER STATISTICS TOOL

USER'S GUIDE

U.S. ENVIRONMENTAL PROTECTION AGENCY
OFFICE OF SOLID WASTE AND EMERGENCY RESPONSE
OFFICE OF SUPERFUND REMEDIATION AND TECHNOLOGY INNOVATION
WASHINGTON, D.C. 20460

September 2018

TABLE OF CONTENTS

Section **Page**

1.0 BACKGROUND/PURPOSE 1
1.1 Remediation Monitoring Phase 1
1.2 Attainment Monitoring Phase 3
2.0 OVERVIEW OF THE GROUNDWATER STATISTICS TOOL 5
2.1 Outlier Testing 5
2.2 Normality Testing 5
2.3 Calculations of the Mean, Linear Trend and Upper Confidence Band 5
2.4 Data Sets with No Detected Values 6
3.0 STEP-BY-STEP INSTRUCTIONS FOR USING THE GROUNDWATER STATISTICS TOOL 7
4.0 EXAMPLES 16
5.0 REFERENCES 19

TABLES

Table **Page**

Table 1. Methods for Calculating UCLs on the Mean by Data Set Type 6
Table 2. Methods for Calculating Linear Trends and Confidence Bands by Data Set Type 6

FIGURES

Figure **Page**

Figure 1. Data Input Screen for Trichloroethene Remediation Example 1, All Eight Data Points Used.....20
Figure 2. Normality Screen for Trichloroethene Remediation Example 1, All Eight Data Points Used.....21
Figure 3. Data Input Screen for Trichloroethene Remediation Example 1, Final Four Data Points Used.....22
Figure 4. Outlier Screen for Trichloroethene Remediation Example 1, Final Four Points Used .23
Figure 5. Normality Screen for Trichloroethene Remediation Example 1, Final Four Points Used.....24
Figure 6. Trend Screen for Trichloroethene Remediation Example 1, Final Four Points Used ..25
Figure 7. UCL Screen for Trichloroethene Remediation Example 1, Final Four Points Used.....26
Figure 8. Site & Summary Statistics Screen for Trichloroethene Remediation Example 1, Final Four Points Used.....27
Figure 9. Data Input Screen for Trichloroethene Attainment Example 2, First Eight Data Points Used.....28
Figure 10. Site & Summary Statistics Screen for Trichloroethene Attainment Example 2, First Eight Data Points Used 29
Figure 11. Data Input Screen for Trichloroethene Attainment Example 2, All Ten Data Points Used.....30
Figure 12. Site & Summary Statistics Screen for Trichloroethene Attainment Example 2, All Ten Data Points Used 31
Figure 13. Data Input Screen for Vinyl Chloride Remediation Example 3.....32
Figure 14. Site & Summary Statistics Screen for Vinyl Chloride Remediation Example 3.....33

1.0 BACKGROUND/PURPOSE

The Groundwater Statistics Tool is designed to help evaluate contaminant of concern (COC) concentrations on a well-by-well basis to determine whether a groundwater restoration remedial action is complete. The tool is designed to support the U.S. Environmental Protection Agency (EPA) memorandum, "Guidance for Evaluating Completion of Groundwater Restoration Remedial Actions" (EPA 2013b, referred to as the *Groundwater Restoration Completion Guidance*), and comports with principles outlined in the "Recommended Approach for Evaluating Completion of Groundwater Restoration Remedial Actions at a Groundwater Monitoring Well" (EPA 2014, referred to as the *Recommended Approach*). Both of these guidance documents should be reviewed before using the tool.

The tool is a Microsoft Excel workbook that is intended to evaluate data for a single COC at a single well. Each Excel worksheet ("screen") is protected to prevent accidental overwriting of formulas. The tool was originally developed in Excel 2010 and is currently designed for Excel 2013; using an older version of Excel or a version using personal computer (PC) emulation in a non-PC environment may not allow use of all of the tool's capabilities. The tool should generally be run separately for each well and each COC being evaluated.

New in This Version

- Summary sheets to show all summary statistics and graphs on one page
- Calculations for dates when cleanup level will be achieved (forecast capability)
- Capability to calculate a UCL even when there are negative trends
- Color-coding and format revisions to make the tool more user-friendly

There are two phases of the groundwater restoration process that the tool is designed to evaluate: the remediation monitoring phase and the attainment monitoring phase.

1.1 Remediation Monitoring Phase

The following text from the *Recommended Approach* discusses the remediation monitoring phase:

> "As discussed in the Groundwater Restoration Completion Guidance, the remediation monitoring phase refers to the phase of the remedy where either active or passive remedial activities are being implemented to reach groundwater cleanup levels selected in a decision document. During this phase, groundwater sampling and monitoring data typically are collected to evaluate contaminant migration and changes in COC concentrations over time. The completion of this phase typically provides stakeholders a decision point for starting data collection and evaluation of the attainment monitoring phase. If an active treatment system is being employed at the site, the completion of this phase may also provide stakeholders with an opportunity to evaluate terminating the system, as appropriate, in the vicinity of the well or wells where groundwater restoration completion is being evaluated. If passive systems are being employed at the site, the data used to make the remediation

phase completion conclusion may also be useful as part of the attainment phase evaluation since active systems are not being employed.

The remediation phase at a monitoring well typically is completed when the data collected and evaluated demonstrate that the groundwater has reached the cleanup levels for all COCs set forth in the record of decision (ROD). It is important to note that at any time during the groundwater remediation, conclusions may be made to remove certain COCs from the monitoring program based on their COC-specific trends or presence in the well. If certain COCs are no longer being evaluated in a well, the rationale for discontinuing monitoring may be used, in conjunction with the current well data, to make the conclusion that all COCs have reached their cleanup levels."

The user should note the following considerations regarding data requirements for each well and COC:

- EPA guidance (2014) recommends a minimum of four data points to evaluate completion of this phase. The tool requires a minimum of four detected results to complete statistical calculations because upper confidence limit (UCL) and trend calculations require at least four detected results to provide sufficient statistical confidence in these environmental settings.

- EPA guidance (2014) provides the user with an option to use either a trend or mean test on the available data set to satisfy the remediation monitoring phase evaluation. In this situation, the selection of the best statistical tool is based on the user's professional judgment.

    - Trend test: A trend line is calculated for this test. It is recommended that the trend test should generally be used for a data set that has a sufficiently steep slope extended over time to give a change in concentration over time with data points that cross below the cleanup level. A less steep change (or asymptotic condition) may not lend itself to a trend test. Once the trend is calculated, an upper confidence band on the trend line should be calculated to allow the user to account for variability within the data set. The use of the upper confidence band on the trend line accounts for uncertainty and provides confidence that the COC cleanup level has been achieved. In general, a 95 percent confidence level is recommended for calculating the upper confidence band.

    - Mean test: It is recommended that the mean test should generally be used for a data set that does not have a sufficiently steep slope extended over time resulting in a dramatic change in concentration. An overly steep slope over an extended period of time results in a change in concentration that would provide high variability in the data set and would tend to elevate the UCL on the mean. Statistics are used to determine the mean contaminant concentration from these data for the COC for this test. The UCL is calculated once the mean is established. The UCL should be compared against the cleanup level. The use of the UCL value accounts for uncertainty and provides confidence that the COC

cleanup level has been achieved. In general, the 95 percent UCL is used as the recommended confidence limit.

1.2 Attainment Monitoring Phase

The following text from the *Recommended Approach* discusses the attainment monitoring phase:

> "The attainment monitoring phase typically occurs after a Region makes a determination that the remediation monitoring phase is complete. When the attainment monitoring phase begins, data typically are collected to first evaluate whether the well has reached steady-state conditions where active remediation activities, if employed, are no longer influencing the groundwater in the well. Once the groundwater is observed to have reached steady-state conditions, data should be collected and evaluated to confirm the attainment monitoring phase has been completed."

The attainment monitoring phase at a monitoring well typically is complete when contaminant-specific data provide a technical and scientific basis that:

(1) The contaminant cleanup level for each COC has been achieved; and

(2) The groundwater will continue to meet the contaminant cleanup level for each COC in the future.

The user should note the following considerations regarding data requirements:

- EPA guidance (2014) recommends a minimum of eight data points generally be used to evaluate completion of this phase. However, the tool requires a minimum of four data points to run the attainment monitoring phase calculations to accommodate site-specific cases. In addition, the tool requires a minimum of four detected results to complete statistical calculations because UCL and trend calculations require at least four detected results to be reliable.
- EPA guidance (2014) recommends evaluating data to determine whether steady-state conditions have been reached for a COC in a well. If steady-state conditions have been reached, it may be appropriate to include some of the remediation monitoring phase data points in the data set used to evaluate the attainment monitoring phase.
- EPA guidance (2014) recommends evaluating the attainment monitoring phase by applying **both** the UCL on the mean and the trend test.
    - UCL on the mean: Statistics are used for this analysis to determine the mean contaminant concentration from the data for the COC. The UCL is calculated once the mean is established. The UCL should be compared with the cleanup level. Using the UCL value accounts for uncertainty and provides confidence that the COC cleanup level has been achieved. In general, the 95 percent UCL is used as the recommended confidence limit.

- Trend: Statistics are used to determine the trend line for this analysis. The slope of the trend line is used to make a conclusion on future groundwater remedial conditions. If the trend line has a zero or negative slope, indicating no or decreasing concentration trends, it may be appropriate to conclude that the contaminant concentrations for each COC in groundwater will remain below the cleanup level. If the trend line has a positive slope that is not statistically significant, then users should consider whether the estimated date when the concentration is predicted to exceed the cleanup level is sufficiently far in the future. If the trend has a statistically significant positive slope, it may be appropriate to evaluate the effectiveness of the remedial action or determine if external contaminant contributions are present.

2.0 OVERVIEW OF THE GROUNDWATER STATISTICS TOOL

The tool is designed to provide a statistical evaluation of a data set to determine that the expectations outlined in EPA guidance (2013b and 2014) and outlined above in this User's Guide have been achieved. In addition to providing statistical mean, trend, and UCL calculations, the tool also tests for detection frequency, outliers, and normality. These tests are described in the following sub-sections. Most of the statistical tests used in the tool are based on the "Unified Guidance for Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities," which will be referred to as the *Unified Guidance* (EPA 2009).

2.1 Outlier Testing

Outliers are checked with a Dixon's test, which is appropriate for data sets composed of fewer than 25 samples. For more details on outliers and the mechanics of the Dixon's test, see Section 12.3 of the *Unified Guidance* (EPA 2009). A significance level of 1 percent is recommended for the Dixon's test and is established as the default in cell B19 of the Data Input screen. The corresponding confidence level is 100 – 1 = 99 percent. However, alternative significance levels for the test can be specified (10 percent, 5 percent, or 0.5 percent). **Dixon's test is used only to indicate whether a data point can be considered as an outlier statistically; outliers should not be discarded from the data set unless there is also a valid, known technical reason for the outlier (for example, field or lab conditions).**

2.2 Normality Testing

The normality of the data set is checked using a Shapiro-Wilk test. For details on normality and the mechanics of the Shapiro-Wilk test, see Section 10.5 of the *Unified Guidance* (EPA 2009). Different levels of significance are used based on the size of the data set (n = number of samples in data set):

- $n < 10$ – the tool uses a significance level of 10 percent
- $10 \leq n < 20$ – the tool uses a significance level of 5 percent
- $n \geq 20$ – the tool uses a significance level of 1 percent

2.3 Calculations of the Mean, Linear Trend and Upper Confidence Band

A confidence level of 95 percent is recommended for the mean calculations and represents the common practice for calculating the UCL on the mean. As shown in Table 1, the UCL on the mean is calculated in different ways, depending on the characteristics of the data set.

A confidence level of 95 percent is recommended to calculate confidence around a trend line and represents the common practice for this calculation. The linear trend and upper confidence band are calculated in different ways, depending on the characteristics of the data set, as shown in Table 2.

Table 1. Methods for Calculating UCLs on the Mean by Data Set Type

Detection Frequency	Type of Data Set (Based on Shapiro-Wilk Test on Data)	
	Normal	Nonparametric
DF = 100%	Student's-t UCL	Chebyshev UCL
0% < DF < 100%	KM Chebyshev UCL	KM Chebyshev UCL
0%	No UCL is calculated	No UCL is calculated

Notes:
KM Kaplan-Meier
DF Detection frequency (percentage of values above the detection limit)
UCL Upper confidence limit

Table 2. Methods for Calculating Linear Trends and Confidence Bands by Data Set Type

Detection Frequency	Type of Data Set (Based on Shapiro-Wilk Test on Residuals)	
	Linear	Nonlinear
DF = 100%	Linear Regression	Theil-Sen line, Mann-Kendall test
0% < DF < 100%	Linear Regression*	Theil-Sen line, Mann-Kendall test
0%	No trend is calculated	No trend is calculated

Notes:
* Before the linear regression is calculated, each nondetect result is substituted with a randomly generated real number between zero and the reported detection limit. This substitution prevents the introduction of artificially low variability from multiple identical (or similar) detection limits.
DF Detection frequency (percentage of values above the detection limit)

2.4 Data Sets with No Detected Values

As noted in Tables 1 and 2, the tool will not calculate a UCL on the mean, a trend line, or an upper confidence band if none of the concentrations are above the detection limit. In this situation, the project team can generally conclude that the phase being evaluated is complete by a visual, or a non-statistical, evaluation if all the detection limits for nondetects in the data set are below the contaminant cleanup level.

3.0 STEP-BY-STEP INSTRUCTIONS FOR USING THE GROUNDWATER STATISTICS TOOL

This tool takes the user through one data input and five results screens. Each screen is protected to prevent accidental overwriting of formulas. These screens are:

1. The "Data_Input" screen.
2. The "Outliers" screen.
3. The "Normality" screen.
4. The "Trend" screen.
5. The "UCL" screen.
6. The "Site & Summary Statistics" screen.

In addition to the data input and results screens, the tool provides a screen titled "Example Data." This screen provides several synthesized data sets that illustrate various functions of the tool. These example data sets are discussed in the Examples section. Example data sets 1 through 3 are from EPA 2013b. Example data sets 4 through 8 can document the tool's performance compared with example results calculated in the *Unified Guidance* (EPA 2009).

Step 1: On the "Data_Input" screen, enter descriptive data into cells B4 to B21, D7 to F26, and J31 to M31. Descriptive data capture information to describe the site, the monitoring phase being evaluated, the COC, and related information that describe the site and purpose of the evaluation. The cells for entering descriptive data are color-coded green; cells in blue are locked and show outputs that the user cannot alter. Control buttons are located in cells that are color-coded red; use these control buttons to move forward and backward through the tool. Follow the instructions below for data entry:

a. Enter the name of the site in cell B4. No restrictions.

b. Enter the operable unit (OU) in cell B5 (if applicable). No restrictions.

c. Select the type of evaluation in cell B6. This value should be either "Remediation" or "Attainment," and there is a drop-down menu to aid in the selection.

d. Enter the date of the evaluation in cell B7. There are no restrictions on how the date is entered. Examples include: March 1, 2014, 3/1/2014, 3/1/14, March 2014.

e. Enter the name of the person performing the evaluation in cell B8. No restrictions.

f. Enter the COC in cell B10. No restrictions.

g. Enter the well name or number in cell B11. No restrictions.

h. Enter the units for date in cell B12. This value should be "Day," "Month," "Year," or "Date," and there is a drop-down menu to aid in the selection. The entries are expected to be dates, or may be days, months, or years from a starting value (which may be zero). The date units should be selected prior to data entry in Columns D through F to ensure proper formatting. For example, if cell B12 is set to some value other than "Date" when

dates have already been entered into the data entry table in column D, the values in the tables and graphs will be displayed as large integer values because of the way Excel handles dates. The beginning value for Excel is that Day 1 is January 1, 1900. For example, May 30, 2018 would be displayed as 43,250.

i. Enter the units for concentration in cell B13. No restrictions. These units should be the same as the units for the cleanup level in cell B16. The preference is to use parts per billion (ppb) or micrograms per liter (µg/L); entering other units may result in issues with significant digits.

j. Enter the confidence level desired for the UCL and trend analyses in cell B15 using the drop-down menu to aid in the selection. The default is the recommended confidence level of 95 percent.

k. Enter the cleanup level for the COC in cell B16. This value should be a positive number and should be extracted from the remedy decision document. Be sure that the units correspond to the entry in cell B13.

l. Enter the type of cleanup level being compared in cell B17/B18 (for example, Maximum Contaminant Level [MCL] or risk-based concentration). No restrictions. The type of cleanup level should be extracted from the remedy decision document.

m. Enter the risk of false outlier rejection in cell B19. This is the significance level for the outlier test. This value should be "0.5%," "1%," "5%," or "10%," and there is a drop-down menu to aid in the selection. The default is 1 percent. Selecting a higher value increases the probability that a data point may be falsely flagged as a potential outlier; selecting a lower value decreases this probability.

n. Random Seed: Cell B20 is a placeholder for a random seed. The random seed is used to calculate imputed values for nondetect data points. The random seed is also used to calculate the time-dependent upper confidence band for nonparametric data sets using a bootstrap procedure. In general, the user should leave this value blank when the data are initially entered. If nondetects are present in the data set or if it is determined that the data set is nonparametric, the tool will automatically generate a random seed. Since the random seed is generated off an internal clock, the value (assuming nothing is entered in this cell during data entry) will be different each time the tool is run. After a data set analysis is conducted, the user should document the random seed value. Documenting this value will allow the results to be replicated. If you desire to replicate results, you may enter this random seed value (a positive real number) in cell B20 for subsequent data calculations.

o. Enter the number of significant figures to use for reporting UCLs and predicted concentrations in cell B21. This value should be a positive whole number greater than zero. The tool restricts the number of significant figures to be between 1 and 3, since analytical precision is not expected to exceed three significant figures. The default value is 2. A drop-down menu is provided to assist with the selection.

p. Time, concentration data and data qualifier may be entered manually into the data entry table in cells D7 to F26 using the keyboard; however, the recommended method is to

enter data by copying and pasting from another Excel spreadsheet or some other compatible electronic data deliverable format to avoid transcription errors. Excel's "paste values" method can be used to avoid overwriting the formatting of the cells.

The tool's data entry table is set up for a maximum of 20 data points. If more than 20 data points are available, it is recommended that only the most recent 20 be used. If the user wants to conduct a trend analysis of more than 20 data points, other software tools can be used (for example, ProUCL [EPA 2013a] or MAROS [AFCEE 2006]).

1. Enter the measurement dates in column D between rows 7 to 26. These dates are expected to be actual dates, or may be days, months, or years from a starting value (which may be zero). A few notes regarding the data set:
    i. The type of date input (day, month, or actual date) should be consistent throughout the data set. This allows for accurate trend analysis.
    ii. All data should be entered in increasing chronological order (oldest to newest).
    iii. If duplicate data are available for a specific date, only one data entry for that date should be made.
2. Enter the measured concentrations in column E between rows 7 and 26. These values are restricted to real numbers greater than zero.
    i. If the date corresponds to a duplicate sampling event, it is recommended that the user enter either the maximum concentration or the average concentration in the concentration field, as agreed by the project team.
    ii. If the measurement is categorized as a nondetect, it is recommended that the user enter the reporting limit for the analytical method. **The user should NOT enter zero or an arbitrary number.**
3. Enter the data qualifier in column F between rows 7 and 26. For purposes of the tool, enter the detection limit for column E as "ND" in column F, even if the laboratory reports results as "U." Based on the qualifier, column G will be populated with a "Yes" or "No" to indicate whether the concentration is a detected result (for a "Yes") or a detection limit for a nondetected result (for a "No"). Blank cells are allowed, and qualifiers should not be entered if they are not applicable.

q. The default selection for the time and concentration axes is "Auto" based the range of data entered in Columns D7 through E26. When "Auto" is entered in lower case, upper case, or mixed case, Excel's automatic setting for time (D7 to D26) and concentration (E7 to E26) will be used for the charts. If the user wants to view the data differently or show only a subset of data, the minimum and maximum values for the date and concentration axes to be displayed may be entered in cells J31 to M31.

r. Click the button labeled "Reset Axes" to adjust the axes in the plots. Adjusting these values can make the data plots easier to view. Once the axes are changed, the tool does not automatically reset the values until the "Reset Axes" button is clicked. The user

can enter "Auto" and click "Reset Axes" to resume Excel's automatic setting for the chart.

Step 2: Review the Data Review and Recommendations section at the bottom of the "Data_Input" screen. If no recommendations appear in red text, the data are ready for statistical analysis. If recommendations appear in red text, the user should follow the recommendations before proceeding with the statistical analysis. Although the statistical analysis can be run without addressing the recommendations, error messages will result.

Step 3: Click the button labeled "Next Step: Check for Outliers." The tool will run Dixon's tests to check for outliers and will display the results of the tests on the "Outliers" screen.

a. Rows 5 and 6 show the number of data points and the selected risk of false rejection, both of which were previously entered on the "Data_Input" screen.

b. Row 7 shows the critical value used for the Dixon's test, which is based on the number of data points and the selected risk of false rejection.

c. Row 8 shows the outlier type being tested. Both low concentration (column B) and high concentration (column C) outliers are tested.

d. Row 9 shows the test statistic. For details on how the test statistic is calculated, refer to Section 12.3 of the *Unified Guidance* (EPA 2009). The test statistic is compared with the critical value to determine whether the lowest concentration data point is a potential low outlier or the highest concentration data point is a potential high outlier. If the test statistic is greater than the critical value, a potential outlier is present.

e. Row 10 shows the result of the test. "Yes" is displayed if a potential outlier is indicated; otherwise "No" is displayed.

f. Row 11 uses the results of the normality testing (which will be shown in row 14 on the next screen) to indicate whether the Dixon's test is valid. See Step 4e below. Since Dixon's test requires that the data are normally distributed, the test is not valid and should not be used to justify an outlier if the remaining data (after excluding any potential outliers) are not normally distributed.

g. If row 10 of the "Outliers" screen indicates no potential outliers, the user may proceed by clicking the button labeled "Next Step: Normality Screen."

h. If row 10 of the "Outliers" screen indicates that potential outliers are present in the data set and row 11 indicates that the Dixon's test is valid for those potential outliers, the user may want to consider revising the data set to remove the outliers. Outliers should be removed only if a valid technical reason for the outlier is known and the project team or user concludes an acceptable reason exists to remove the outlier. There are two options for proceeding in the presence of an outlier:

    i. Elect to remove the outlier. The outlier can be removed from the data set in one of two ways: 1) click the button labeled "Previous Step: Data Input Screen" to return to the "Data Input" screen or 2) click the button labeled "Next Step:

Normality Screen", where a pop up box will be displayed. To get back to the "Data Input" screen, an "N" should be entered in the prompt.

Upon returning to the "Data Input" screen to remove the outlier(s) from the data set, the user should first delete the "date" (column D) and "concentration" (column E) values for the outlier(s). The user should then "Copy" the date and concentration values for all data that was collected chronologically after the outlier(s). The user should then paste the copied cells into the table starting in the row where the original outlier(s) was deleted. The user should then delete the data in the final row(s) to ensure there is no duplicative data in the set (however, the entire row should not be deleted).

Note: When modifying the data set, do not use the "Cut" function. Instead, use "Copy and Paste" to move data and the "Clear contents" function to remove values from the data set.

ii. Retain full data set. If the user elects to retain the outlier(s) and move forward with the statistical analysis, the user should click the button labeled "Next Step: Normality Screen" and enter "Y" into the prompt field.

Step 4: The tool will run three Shapiro-Wilk tests on the data set to determine whether the data and the residuals appear to be normally distributed. The three test results are shown on the "Normality" screen. **The first column is for the full data set and is used to determine which UCL calculation method (normal or nonparametric) should be used.** The second column is for the data set minus the potential outliers and is used only to determine whether the Dixon's test is valid. Dixon's test is not considered valid if the data set minus the potential outliers does not appear to be normal. If no outliers exist, the column for the data set minus potential outliers will not be populated with results**. The third column is for the residuals and is used to test whether parametric or nonparametric methods are used for fitting the trend.** The following results are shown for each test:

a. Rows 7 and 8 show the number of data points and the selected alpha values for the Shapiro-Wilk test. The number of data points was previously calculated in cell B23 on the "Data_Input" screen. The alpha value for the Shapiro-Wilk test is chosen by the tool based on the number of data points (see section 2.2 above).

b. Rows 9, 10 and 11 show the slopes, intercepts and correlation coefficients (R) of the normal Q-Q plot. These are shown for informational purposes only.

c. Row 12 shows the exact Shapiro-Wilk test values, which are the test statistics. For more detail on how the test value is calculated, see Section 10.5 of the *Unified Guidance* (EPA 2009).

d. Row 13 shows the critical values for each test, which is based on the number of data points and the Shapiro-Wilk alpha values. For more detail on how the critical value is determined, see Section 10.5 of the *Unified Guidance* (EPA 2009).

e. Row 14 shows the result of the tests. If the exact test value is lower than the critical value, the conclusion of the test is that the data (or the residuals, for the third column) do not appear normal. If the exact test value is equal to or greater than the critical value, the conclusion of the test is that the data (or the residuals, for the third column) appear normal.

The user will be shown a message summarizing the results of the normality tests and the analysis methods to be used for the UCL and trend calculations. The user should press the OK button before proceeding.

The user may now proceed to either the "Trend" screen or the "UCL" screen, or return to the "Outliers" screen by using the buttons at the bottom of the screen.

If the "Trend" screen is selected, then an error message will appear if the trend predicts negative concentrations during the time period of the measurements entered into the tool. In this case, the message will inform the user that the trend in the data set cannot be evaluated by the tool and the "Trend" screen will not be shown. This situation can occur when there is a steep trend in the data that ends (or starts) near a concentration of zero. This situation is not expected to occur commonly; if it does occur, there are two ways to correct the situation.

1. Data points can be removed from the data set. Often, a case may be observed where a group of data points within the data set has a steeply decreasing trend and these data points are followed by a group of data points with a less steep trend. In this case, the early points where the concentration is decreasing rapidly should be removed, while the number of data points recommended for the decision (as described in Section 1) is maintained. Data points at the end of the data set should not be removed.
2. Normalizing transformations of the data may be attempted to account for the change in slope. This tool (or another tool such as ProUCL [EPA 2013a]) may still be used for trend testing, but the data must be transformed outside the tool and imported into the tool. The user may want to consult a statistician for assistance with this process. In this situation, the UCL must be calculated on the untransformed data; only the trend should be calculated using the transformed data. For more details on other transformations that can be attempted with the data, see Section 3.2.4 of the *Unified Guidance* (EPA 2009).

Step 5: Inspect the results on the "Trend" screen.

- If the trend is found to be linear, the trend will be evaluated using the ordinary least squares method. The following outputs will be shown:
    - The data set will be shown in columns B and C of the table. The predicted concentrations from the ordinary least squares method, the residuals (difference between the measured and predicted concentrations), and the calculated values of the upper confidence band for each data point will also be shown in columns D, E, and F of the table. The confidence level for the upper confidence band is the (two-sided) confidence level that was selected in cell B15 of the "Data_Input"

screen. For example, if the selected confidence level is the default 95%, then the upper confidence band is the estimated 97.5th percentile concentration.

- The slope, intercept and R-squared values of the linear regression trend line will be shown in cells I7, I8, and I9.
- The test statistic will be shown in cell I11. For details on calculation of the test statistic, see Section 17.3.1 of the *Unified Guidance* (EPA 2009). The test statistic has the same sign as the slope.
- The critical value of the test will be shown in cell I12. The critical value is based on the number of data points and the confidence level selected on the "Data_ Input" screen and will have a positive sign.
- The test result will be shown in cell I10. If the test statistic is positive and the test statistic is greater than the critical value, the conclusion is that the trend is increasing. If the test statistic is negative and the absolute value of the test statistic is greater than the critical value, the conclusion is that the trend is decreasing. Otherwise, the conclusion is that there is no trend. The color of this cell corresponds to the value of the test result: red is "Increasing," green is "Decreasing" and yellow is "No trend."
- The Trend Type for this case is "Normal," as shown in cell I13.
- If the slope is negative and the concentration is predicted to achieve the cleanup level on some future date *after* the date of the last measurement, then that future date is shown in cell I14. This date is shown even if the trend is not statistically significant. Otherwise the value in cell I14 is shown as "NA." The date is shown as "NA" if the concentration is predicted to have already achieved the cleanup level.
- If the slope is positive and the concentration is predicted to reach the cleanup level on some future date *after* the date of the last measurement, then that future date is shown in cell I15. This date is shown even if the trend is not statistically significant. Otherwise the value in cell I15 is shown as "NA." The date is shown as "NA" if the concentration is predicted to have already reached the cleanup level.

- If the trend is not found to be linear, the trend will be evaluated using the Mann-Kendal test using a Theil-Sen slope. The following outputs will be shown:
  - The data set will be shown in columns B and C of the table. The predicted concentrations from the Theil-Sen nonparametric trend line, the residuals (difference between the measured and predicted concentrations), and the calculated values of the upper confidence band for each data point will also be shown in columns D, E, and F of the table. The confidence level for the upper confidence band is the (two-sided) confidence level that was selected in cell B15 of the "Data_Input" screen. For example, if the selected confidence level is the

default 95%, then the upper confidence band is the estimated 97.5th percentile concentration.

- The critical value of the Mann-Kendall test will be shown in cell I10. The critical value is based on the confidence level selected on the "Data Input" screen and will have a positive sign.
- The test statistic (S) and normalized value will be shown in cells I8 and I9. For details on the calculation of these values, see Section 17.3.2 of the *Unified Guidance* (EPA 2009). The test statistic has the same sign as the slope. The test statistic in I8 is a measure of whether or not there is a trend in the data.
- The slope and intercept of the Theil-Sen trend line will be shown in cells L7 and L8.
- The test result will be shown in cell I7. If the test statistic (S) is positive and the normalized test statistic is greater than the critical value, the conclusion is that the trend is increasing. If the test statistic is negative and the absolute value of the normalized test statistic is greater than the critical value, the conclusion is that the trend is decreasing. Otherwise, the conclusion is that there is no trend. The color of this cell corresponds to the value of the test result: red "Increasing," green is "Decreasing" and yellow is "No trend."
- The Trend Type for this case is "Nonparametric," as shown in cell I11.
- If the slope is negative and the concentration is predicted to achieve the cleanup level on some future date *after* the date of the last measurement, then that future date is shown in cell L9. This date is shown even if the trend is not statistically significant. Otherwise the value in cell L9 is shown as "NA." The date is shown as "NA" if the concentration is predicted to have already achieved the cleanup level.
- If the slope is positive and the concentration is predicted to reach the cleanup level on some future date *after* the date of the last measurement, then that future date is shown in cell L10. This date is shown even if the trend is not statistically significant. Otherwise the value in cell L10 is shown as "NA." The date is shown as "NA" if the concentration is predicted to have already reached the cleanup level.

- The trend line is displayed on a plot. The confidence band on the trend will also be displayed. To see the future trends on the trend chart, select Axis Values in cells B34 to E34 and click the "Reset Axes" button.

The user may now proceed to the "UCL" screen.

Step 6: Inspect the results on the "UCL" screen. The screen shows the cleanup level, the estimated mean, the estimated standard deviation, and the calculated upper confidence limit for the mean (UCL). It also includes a summary of the data entered on the "Data_Input" screen. If nondetects are present, the tool displays a table that summarizes the data and the imputed

values assigned by the tool. The tool displays a graph that shows the UCL, mean, cleanup level, and the data values.

The user may now proceed to the "Site & Summary Statistics" screen.

Step 7: Inspect the results on the "Site & Summary Statistics" screen. Note that the name and format of this tab varies depending upon the normality or non-normality of the data and residuals, whether or not there are negative predictions, and whether or not there are nondetects. The screen shows a compilation of the data analysis and results presented in preceding sheets, including the Trend and UCL graphs. The results are summarized in three panels. The "General Information" panel is a summary of the data inputs from the "Data_Input" screen. The "Trend Analysis" panel includes the trend type and analysis method, a Yes/No indicator for whether the maximum upper bound is greater than the cleanup level (for the Remediation phase only; result is NA for the Attainment phase), estimated intercept and slope, the result of the statistical trend test, and the future dates when the concentration is predicted to achieve or exceed the cleanup level, if applicable. Details of the trend analysis are not provided if there are any negative predictions. The "UCL Analysis" panel includes the distribution type and analysis method, mean, UCL, and a Yes/No indicator for whether the UCL is greater than the cleanup level.

4.0 EXAMPLES

As noted earlier, several example data sets are provided on the "Example Data" screen. Each data set is discussed briefly in the sections below. The last seven examples are not meant specifically to test the remediation and attainment decisions, although a decision will be shown on the "UCL" screens based on the settings for the decision type, cleanup level, and confidence levels. For Examples 4 through 8 the corresponding results in the *Unified Guidance* (EPA 2009, Example 12-3, modified Example 12-3, Example 10-2, Example 17-5, and Example 21-7, respectively) may differ slightly from the results calculated by the tool because of differences in rounding. Figures at the end of this guide show screenshots for Examples 1 through 3.

- Example Data Set 1. This data set is for trichloroethene and consists of eight data points. These data are for the remediation monitoring phase at a 95 percent confidence level. There are two sub-examples here:
    - If all eight points are used, the data set is not normally distributed but the residuals are normally distributed; therefore, the 95 percent UCL will be calculated using nonparametric methods while the trends will be calculated using linear regression. The fit to the data results in predictions of negative concentrations, and an error message is generated when the user clicks the "Next Step: Trend Screen" on the "Normality" screen. Screen shots of the "Data_Input" and "Normality" screens for this case are shown in Figures 1 and 2.
    - If only the last four data points are used, where the data are closer to the MCL and more linear, the 95 percent upper confidence band at the final sampling event will be below the MCL, and no negative concentrations are predicted. Although the overall 95 percent UCL of the data set is above the MCL, as shown in Figures 7 and 8, only one criterion must be met to proceed to attainment. This is because the earlier values are well above the cleanup level. Screen shots of the "Data_Input," "Outlier," "Normality," "Trend," "UCL," and "Site & Summary Statistics" screens for this case are shown in Figures 3 through 8.
- Example Data Set 2. This data set is a continuation of the previous data set for trichloroethene and consists of 10 data points for the attainment monitoring phase. Again, there are two sub-examples here:
    - If only the first eight points are used, the data set is found to have a statistically significant decreasing trend, but the calculated 95 percent UCL is not below the MCL. Screen shots of the "Data_Input" and "Site & Summary Statistics" screens for this case are shown in Figures 9 and 10.
    - If all 10 data points are used, the data set is found to satisfy both attainment criteria identified in Section 1.2; the trend is statistically significant and decreasing, and the 95 percent UCL is below the MCL. In fact, this scenario is also the case after nine sampling events. Screen shots of the "Data_Input" and "Site & Summary Statistics" screens for this case are shown in Figures 11 and 12.

- Example Data Set 3. This data set is for vinyl chloride at the same hypothetical site as for Examples 1 and 2 and is for the remediation monitoring phase. The concentration is observed to be increasing toward the MCL and there is a slight elevated result at the seventh event. Although the upper confidence band exceeds the MCL of 2 ppb, the 95 percent UCL does not. Only one criterion must be met to proceed to attainment, so the remediation monitoring phase is considered complete. Screen shots of the "Data_Input" and "Site & Summary Statistics" screens for this case are shown in Figures 13 and 14.
- Example Data Set 4. This data set is Example 12-3 from Section 12.3 of the *Unified Guidance* (EPA 2009). The data are the logged concentrations from the example. There are 20 data points, and the 20th data point is an outlier at the 95 percent confidence level. Additional calculations can be examined for this example, but it is intended to illustrate the tool's use in checking the outlier calculations.
- Example Data Set 5. This data set is modified from the previous data set by subtracting each concentration from 10 ppb. There are 20 data points, and the 20th data point is an outlier at the 95 percent confidence level. Additional calculations can be examined for this example, but it is intended for use in checking the outlier calculations.
- Example Data Set 6. This data set is from Example 10-2 in Section 10.5.1 of the *Unified Guidance* (EPA 2009). The data set has 20 data points and is not normal at the 99 percent confidence level. The example can be checked by proceeding to the "Normality" screen and comparing the test results against the example in the *Unified Guidance*. Additional calculations can be examined for this example, but it is intended for use in checking the normality calculations. If trend calculations are attempted, this example will generate an error message because negative concentrations are predicted for the early data points.
- Example Data Set 7. This data set is from Example 17-5 in Section 17.3.1 of the *Unified Guidance* (EPA 2009). The data set has 19 data points with normal residuals. It can be used to illustrate the linear regression and compare it against the results for the example in the *Unified Guidance*. The trend is shown to be statistically increasing at the 99 percent confidence level in the example.
- Example Data Set 8. This data set is from Example 21-7 in Section 21.3.1 of the *Unified Guidance* (EPA 2009). The data set has 10 data points with normal residuals. It can be used to illustrate both the linear regression and the upper confidence band calculations and compare them against the example. The upper confidence band is calculated at the 95 percent confidence level in the example.
- Example Data Set 9. This data set is the same as example data set 9, described above, with one change. One additional result is added, a nondetect at a detection limit of 2 ppb. This data set can be used to illustrate UCLs calculated based on the nonparametric Kaplan-Meier (KM) Chebyshev UCL. In this example, the residuals are still normal, so the trend will be calculated using linear regression but using an imputed value for the nondetect result. If a system-generated random seed predicts negative concentrations, use a random seed of 62311.421888.

- Example Data Set 10. This data set is based on Example 21-8 in Section 21.3.2 of the *Unified Guidance* (EPA 2009). The data set has 10 data points with normal residuals; the final data point has been altered from the original example so that the trend is not statistically significant at the 95 percent confidence level; this data set illustrates how the output will appear if the trend is not found to be statistically significant.

5.0 REFERENCES

Air Force Center for Environmental Excellence (AFCEE). 2006. Monitoring and Remediation Optimization Systems (MAROS) Software Version 2.2 User Guide. March.

U.S. Environmental Protection Agency (EPA). 2009. Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities – Unified Guidance. EPA 530/R-09-007. March. Accessed on line in January 2014 at: http://www.epa.gov/osw/hazard/correctiveaction/resources/guidance/sitechar/gwstats/

EPA. 2013a. ProUCL Version 5.0.00 Technical Guide, Statistical Software for Environmental Applications for Data Sets with and without Nondetect Observations. September.

EPA. 2013b. Guidance for Evaluating Completion of Groundwater Restoration Remedial Actions. OSWER 9355.0-129. From: James E. Woolford, Director, Office of Superfund Remediation and Technology Innovation; and Reggie Cheatham, Director, Federal Facilities Restoration and Reuse Office. To: Superfund National Policy Managers, Regions 1 - 10; and Federal Facility Leadership Council. November 25.

EPA. 2014. Recommended Approach for Evaluating Completion of Groundwater Restoration Remedial Actions at a Groundwater Monitoring Well. OSWER 9283.1-44. August.

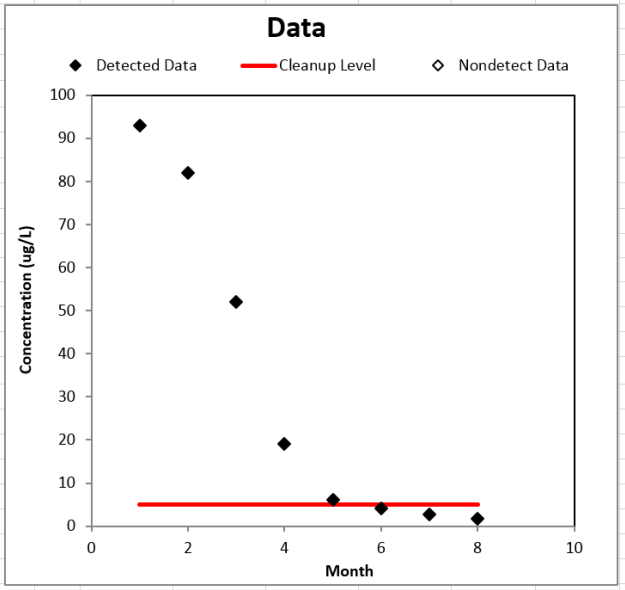
Figure 1. Data Input Screen for Trichloroethene Remediation Example 1, All Eight Data Points Used

Groundwater Statistics Tool

Data input worksheet

Site Name	Test
Operable Unit (OU)	Test
Type of Evaluation	Remediation
Date of Evaluation	4/9/2018
Person performing analysis	DW
Chemical of Concern	TCE
Well Name/Number	1
Date Units	Month
Concentration Units	ug/L
Confidence Level Desired	95%
Cleanup Level	5
Source of cleanup level (e.g. MCL or risk-based concentration)	MCL
Risk of False Outlier Rejection	1%
Random Seed (may be left blank)	62311.42188
Significant figures to use	3
Number of data points:	8
Number of detected results:	8
Number of nondetect results:	0
Detection frequency:	100%

Date (Month)	TCE Conc. (ug/L)	Data Qualifier	Detected? (Yes or No)
1	93		Yes
2	82		Yes
3	52		Yes
4	19		Yes
5	6.1		Yes
6	4.2		Yes
7	2.8		Yes
8	1.8		Yes



Data Review		Recommendations
Are all necessary data fields entered, and in proper format?	Yes	None
Are at least 4 data points present for statistical analysis?	Yes	None
Are at least 8 data points present for an attainment evaluation?	Yes	None
Are detection limits for nondetects ≤ maximum detected value?	Yes	None
Are all data within chart axis limits?	Yes	None

Select Axis Values and Press "Reset Axes" after entering data.

Axis Values			
Time		Concentration	
Min	Max	Min	Max
Auto	Auto	Auto	Auto

Reset Axes

Pressing the "Check for Outliers" button to the right will open a worksheet that shows the results of a Dixons's test for outliers.

Next Step: Check for Outliers

Figure 2. Normality Screen for Trichloroethene Remediation Example 1, All Eight Data Points Used

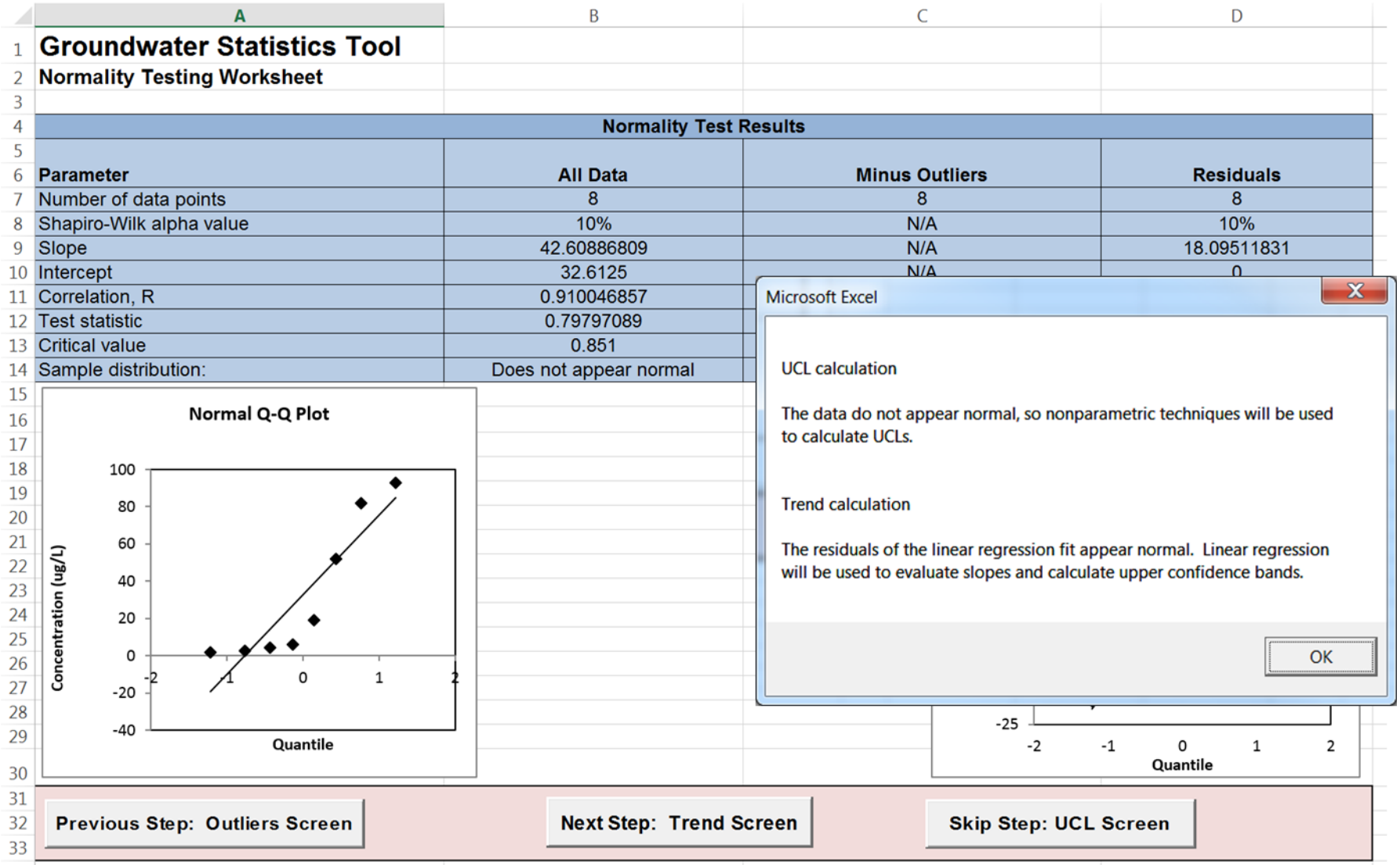
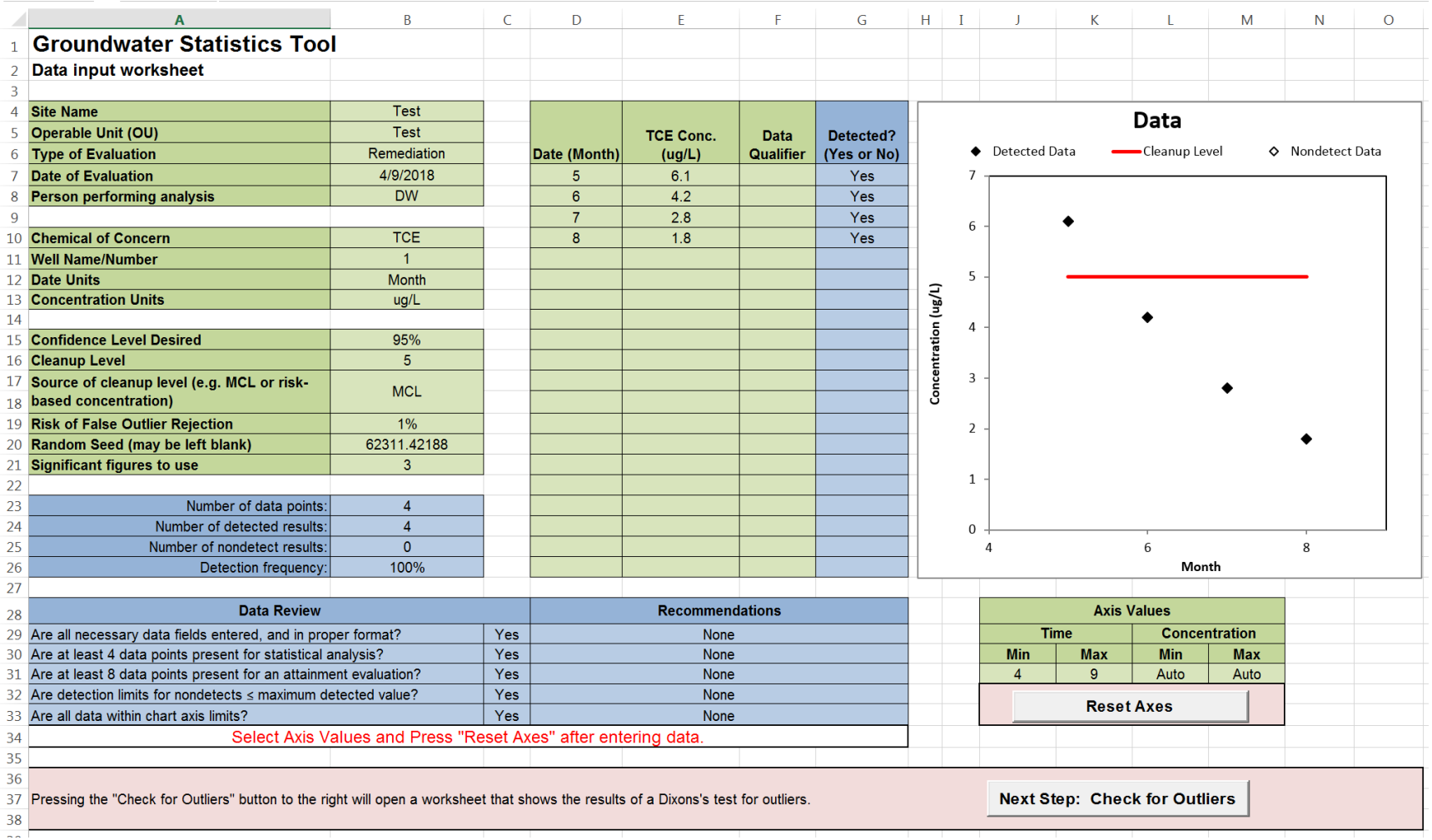


Figure 3. Data Input Screen for Trichloroethene Remediation Example 1, Final Four Data Points Used



Groundwater Statistics Tool

Data input worksheet

Site Name	Test
Operable Unit (OU)	Test
Type of Evaluation	Remediation
Date of Evaluation	4/9/2018
Person performing analysis	DW
Chemical of Concern	TCE
Well Name/Number	1
Date Units	Month
Concentration Units	ug/L
Confidence Level Desired	95%
Cleanup Level	5
Source of cleanup level (e.g. MCL or risk-based concentration)	MCL
Risk of False Outlier Rejection	1%
Random Seed (may be left blank)	62311.42188
Significant figures to use	3
Number of data points:	4
Number of detected results:	4
Number of nondetect results:	0
Detection frequency:	100%

Date (Month)	TCE Conc. (ug/L)	Data Qualifier	Detected? (Yes or No)
5	6.1		Yes
6	4.2		Yes
7	2.8		Yes
8	1.8		Yes

Data Review		Recommendations
Are all necessary data fields entered, and in proper format?	Yes	None
Are at least 4 data points present for statistical analysis?	Yes	None
Are at least 8 data points present for an attainment evaluation?	Yes	None
Are detection limits for nondetects ≤ maximum detected value?	Yes	None
Are all data within chart axis limits?	Yes	None

Select Axis Values and Press "Reset Axes" after entering data.

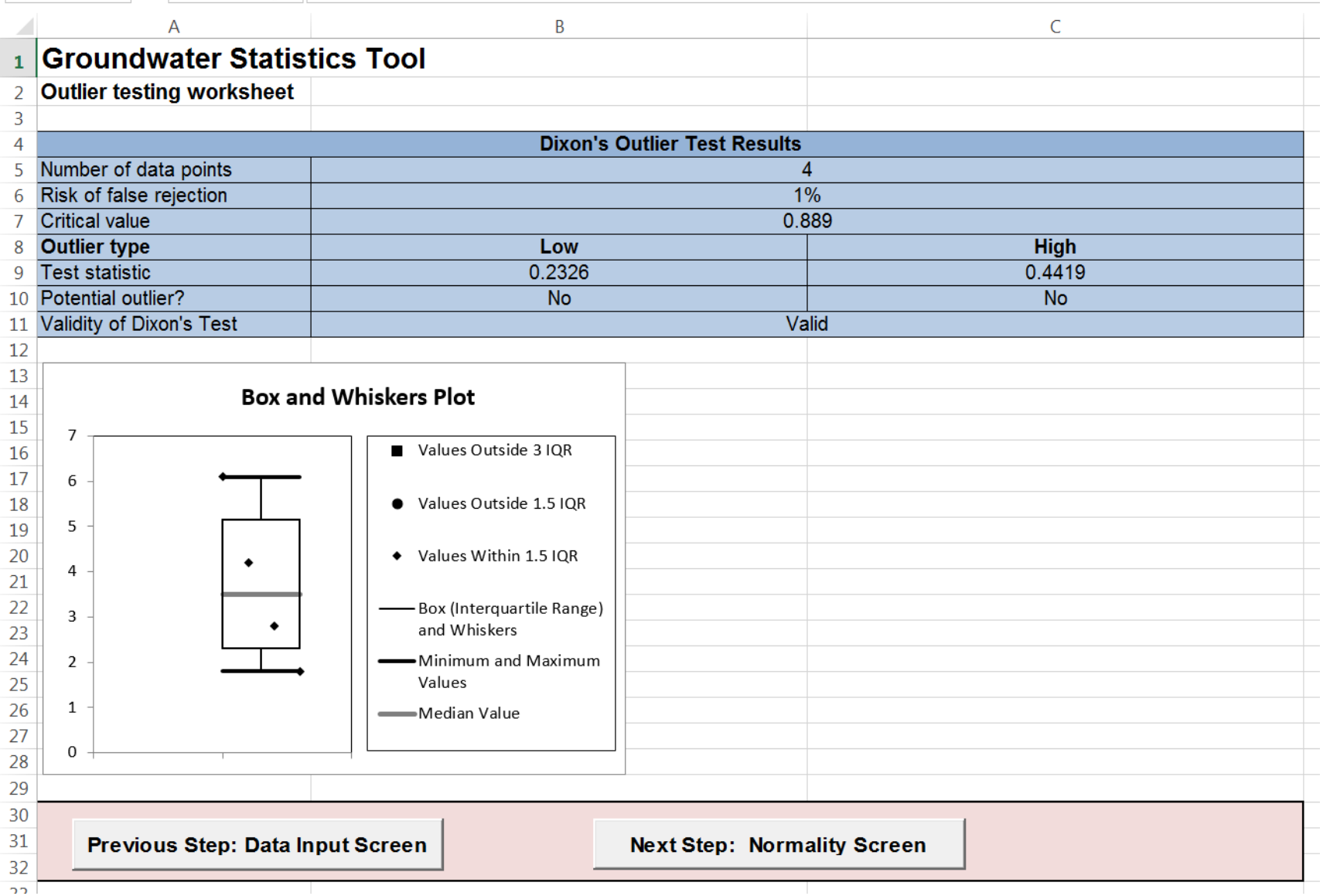
Axis Values			
Time		Concentration	
Min	Max	Min	Max
4	9	Auto	Auto

Reset Axes

Pressing the "Check for Outliers" button to the right will open a worksheet that shows the results of a Dixons's test for outliers.

Next Step: Check for Outliers

Figure 4. Outlier Screen for Trichloroethene Remediation Example 1, Final Four Points Used



Groundwater Statistics Tool

Outlier testing worksheet

Dixon's Outlier Test Results		
Number of data points	4	
Risk of false rejection	1%	
Critical value	0.889	
Outlier type	**Low**	**High**
Test statistic	0.2326	0.4419
Potential outlier?	No	No
Validity of Dixon's Test	Valid	

Previous Step: Data Input Screen

Next Step: Normality Screen

Figure 5. Normality Screen for Trichloroethene Remediation Example 1, Final Four Points Used

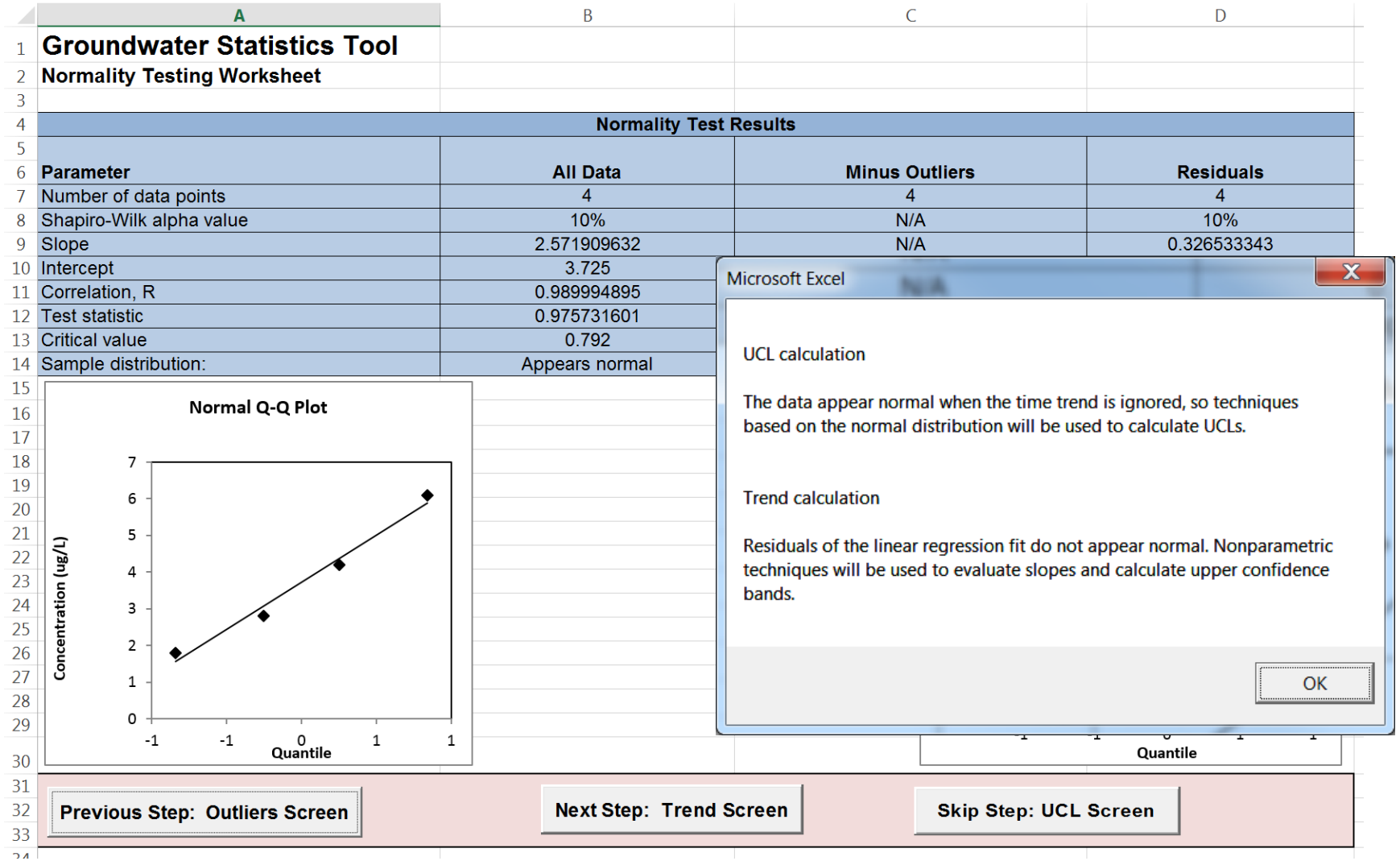


Figure 6. Trend Screen for Trichloroethene Remediation Example 1, Final Four Points Used

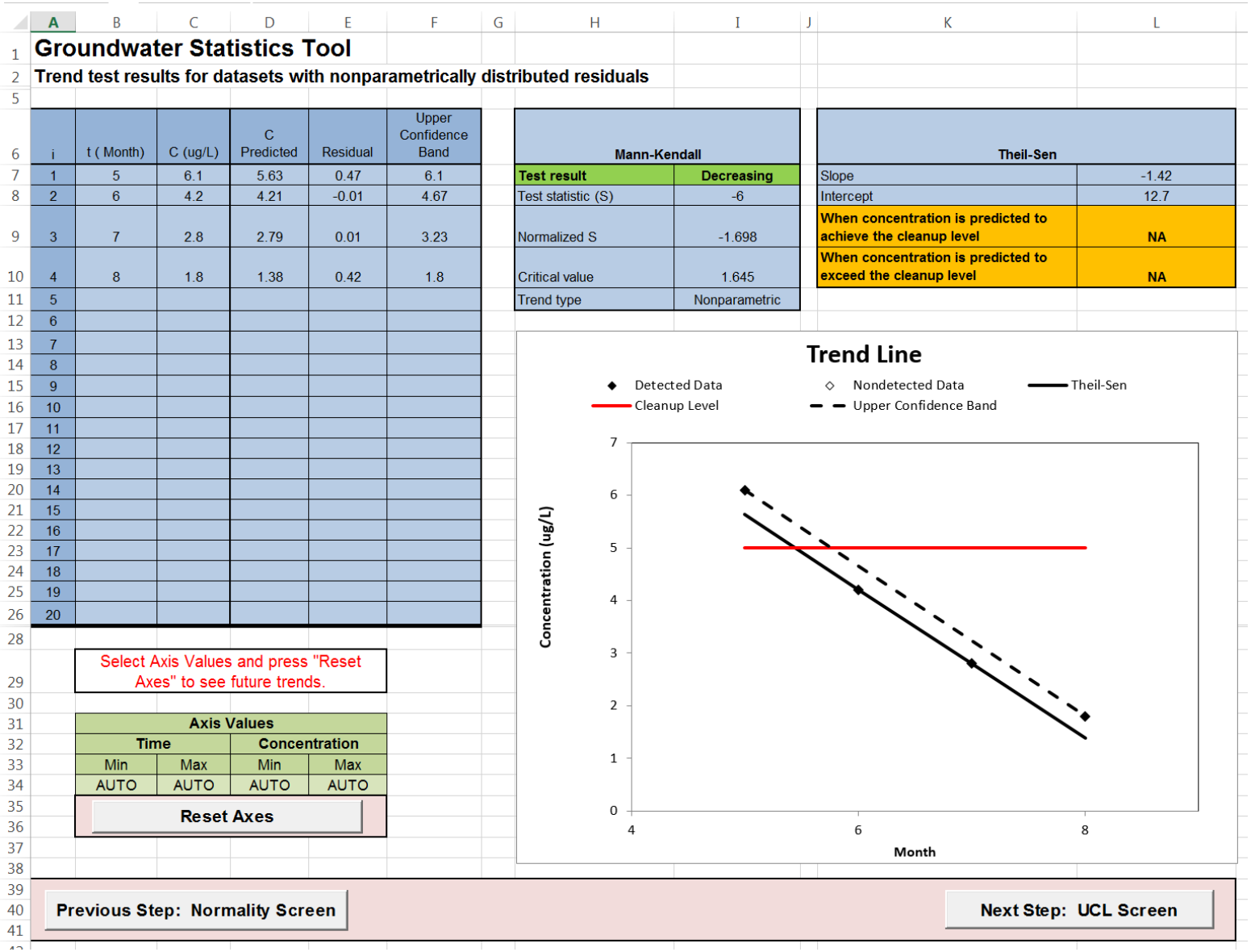


Figure 7. UCL Screen for Trichloroethene Remediation Example 1, Final Four Points Used

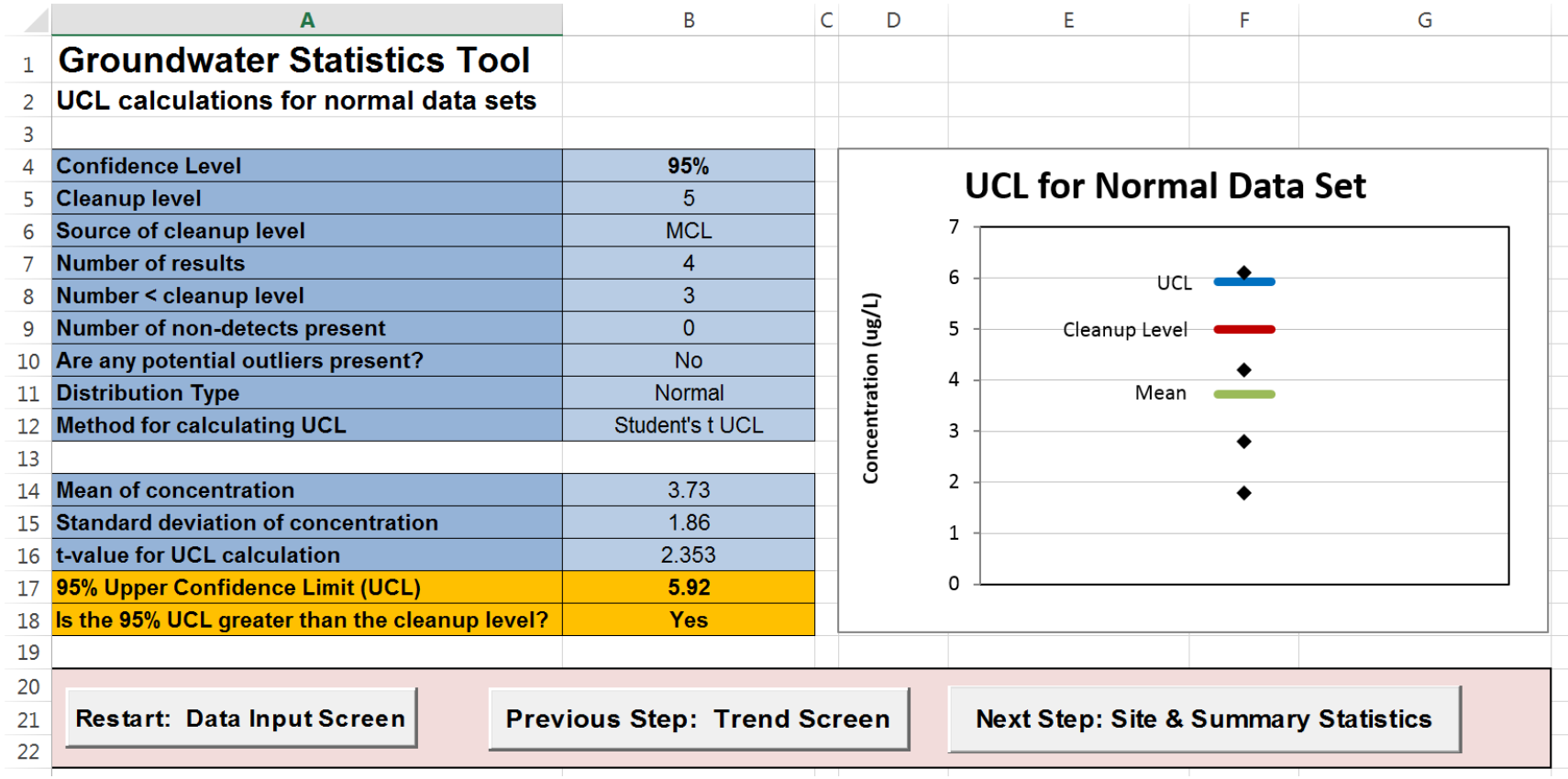
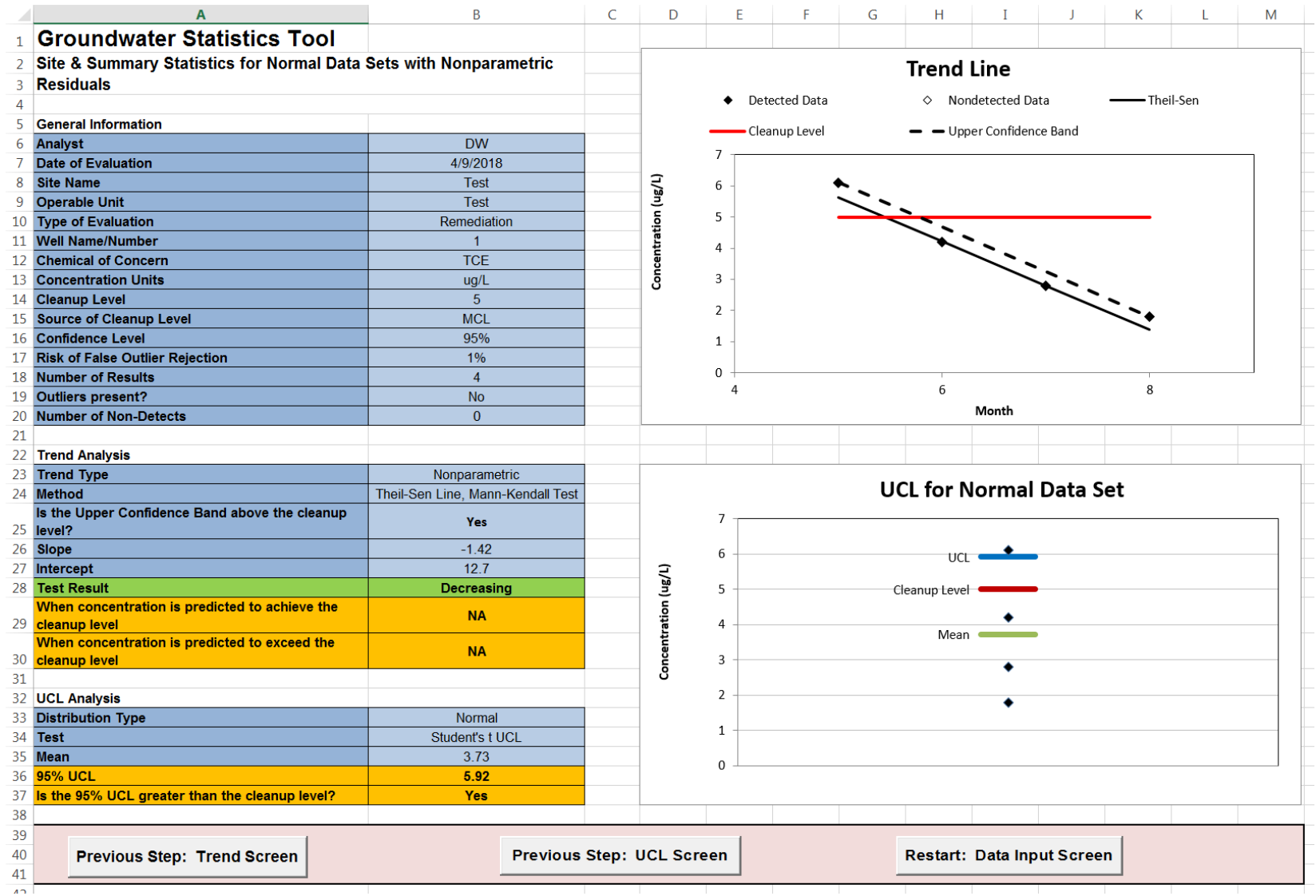


Figure 8. Site & Summary Statistics Screen for Trichloroethene Remediation Example 1, Final Four Points Used



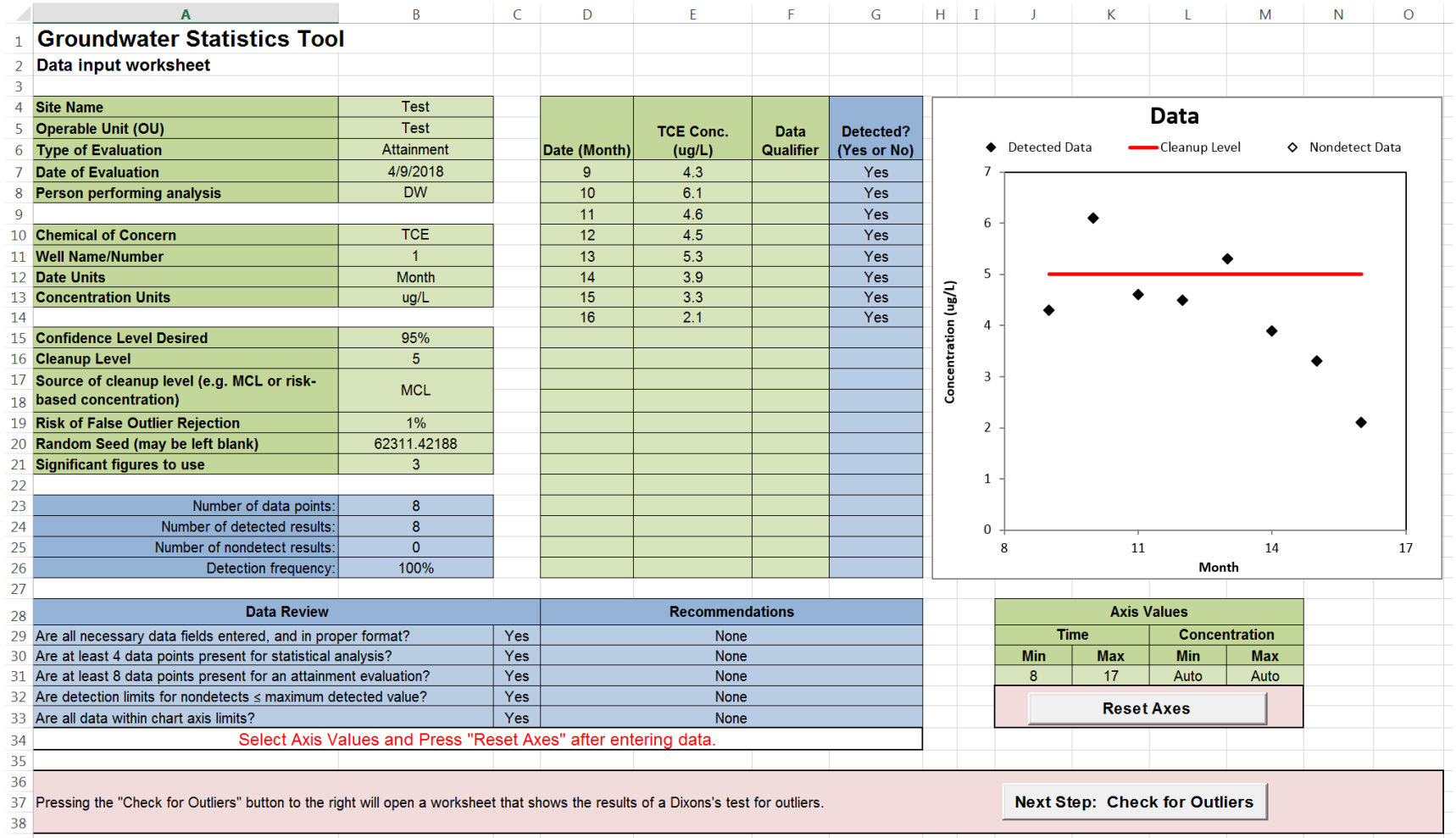
Groundwater Statistics Tool

Site & Summary Statistics for Normal Data Sets with Nonparametric Residuals

General Information	
Analyst	DW
Date of Evaluation	4/9/2018
Site Name	Test
Operable Unit	Test
Type of Evaluation	Remediation
Well Name/Number	1
Chemical of Concern	TCE
Concentration Units	ug/L
Cleanup Level	5
Source of Cleanup Level	MCL
Confidence Level	95%
Risk of False Outlier Rejection	1%
Number of Results	4
Outliers present?	No
Number of Non-Detects	0
Trend Analysis	
Trend Type	Nonparametric
Method	Theil-Sen Line, Mann-Kendall Test
Is the Upper Confidence Band above the cleanup level?	**Yes**
Slope	-1.42
Intercept	12.7
Test Result	**Decreasing**
When concentration is predicted to achieve the cleanup level	**NA**
When concentration is predicted to exceed the cleanup level	**NA**
UCL Analysis	
Distribution Type	Normal
Test	Student's t UCL
Mean	3.73
95% UCL	**5.92**
Is the 95% UCL greater than the cleanup level?	**Yes**

Previous Step: Trend Screen | Previous Step: UCL Screen | Restart: Data Input Screen

Figure 9. Data Input Screen for Trichloroethene Attainment Example 2, First Eight Data Points Used



Groundwater Statistics Tool

Data input worksheet

Site Name	Test
Operable Unit (OU)	Test
Type of Evaluation	Attainment
Date of Evaluation	4/9/2018
Person performing analysis	DW
Chemical of Concern	TCE
Well Name/Number	1
Date Units	Month
Concentration Units	ug/L
Confidence Level Desired	95%
Cleanup Level	5
Source of cleanup level (e.g. MCL or risk-based concentration)	MCL
Risk of False Outlier Rejection	1%
Random Seed (may be left blank)	62311.42188
Significant figures to use	3
Number of data points:	8
Number of detected results:	8
Number of nondetect results:	0
Detection frequency:	100%

Date (Month)	TCE Conc. (ug/L)	Data Qualifier	Detected? (Yes or No)
9	4.3		Yes
10	6.1		Yes
11	4.6		Yes
12	4.5		Yes
13	5.3		Yes
14	3.9		Yes
15	3.3		Yes
16	2.1		Yes

Data Review		Recommendations
Are all necessary data fields entered, and in proper format?	Yes	None
Are at least 4 data points present for statistical analysis?	Yes	None
Are at least 8 data points present for an attainment evaluation?	Yes	None
Are detection limits for nondetects ≤ maximum detected value?	Yes	None
Are all data within chart axis limits?	Yes	None

Select Axis Values and Press "Reset Axes" after entering data.

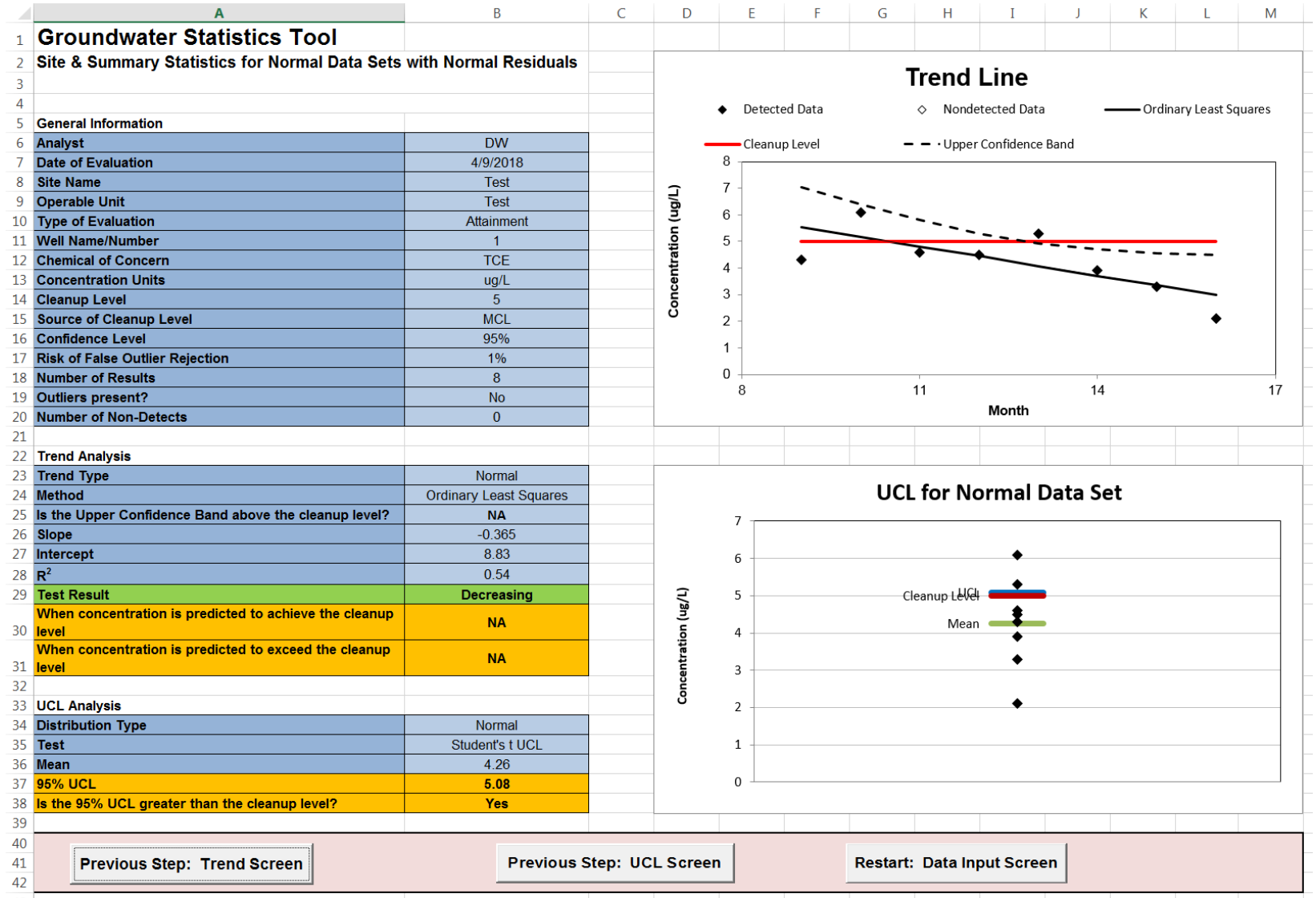
Axis Values			
Time		Concentration	
Min	Max	Min	Max
8	17	Auto	Auto

Reset Axes

Pressing the "Check for Outliers" button to the right will open a worksheet that shows the results of a Dixons's test for outliers.

Next Step: Check for Outliers

Figure 10. Site & Summary Statistics Screen for Trichloroethene Attainment Example 2, First Eight Data Points Used



Groundwater Statistics Tool

Site & Summary Statistics for Normal Data Sets with Normal Residuals

General Information	
Analyst	DW
Date of Evaluation	4/9/2018
Site Name	Test
Operable Unit	Test
Type of Evaluation	Attainment
Well Name/Number	1
Chemical of Concern	TCE
Concentration Units	ug/L
Cleanup Level	5
Source of Cleanup Level	MCL
Confidence Level	95%
Risk of False Outlier Rejection	1%
Number of Results	8
Outliers present?	No
Number of Non-Detects	0
Trend Analysis	
Trend Type	Normal
Method	Ordinary Least Squares
Is the Upper Confidence Band above the cleanup level?	**NA**
Slope	-0.365
Intercept	8.83
R^2	0.54
Test Result	**Decreasing**
When concentration is predicted to achieve the cleanup level	**NA**
When concentration is predicted to exceed the cleanup level	**NA**
UCL Analysis	
Distribution Type	Normal
Test	Student's t UCL
Mean	4.26
95% UCL	**5.08**
Is the 95% UCL greater than the cleanup level?	**Yes**

Previous Step: Trend Screen | Previous Step: UCL Screen | Restart: Data Input Screen

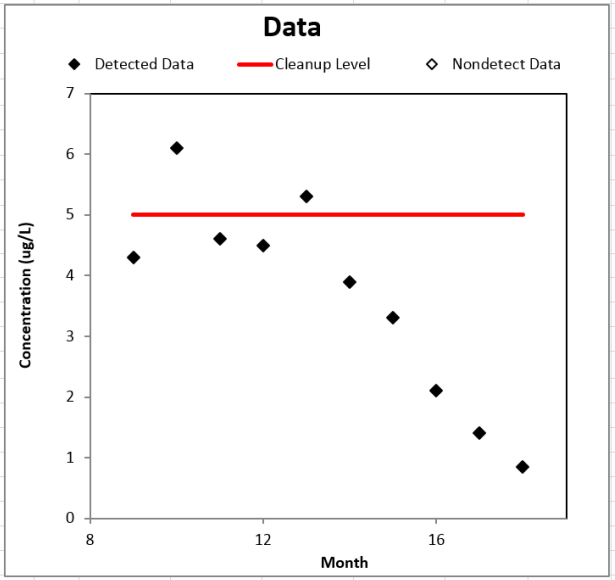
Figure 11. Data Input Screen for Trichloroethene Attainment Example 2, All Ten Data Points Used

Groundwater Statistics Tool

Data input worksheet

Site Name	Test
Operable Unit (OU)	Test
Type of Evaluation	Attainment
Date of Evaluation	4/9/2018
Person performing analysis	DW
Chemical of Concern	TCE
Well Name/Number	1
Date Units	Month
Concentration Units	ug/L
Confidence Level Desired	95%
Cleanup Level	5
Source of cleanup level (e.g. MCL or risk-based concentration)	MCL
Risk of False Outlier Rejection	1%
Random Seed (may be left blank)	62311.42188
Significant figures to use	3
Number of data points:	10
Number of detected results:	10
Number of nondetect results:	0
Detection frequency:	100%

Date (Month)	TCE Conc. (ug/L)	Data Qualifier	Detected? (Yes or No)
9	4.3		Yes
10	6.1		Yes
11	4.6		Yes
12	4.5		Yes
13	5.3		Yes
14	3.9		Yes
15	3.3		Yes
16	2.1		Yes
17	1.4		Yes
18	0.85		Yes



Data Review		Recommendations
Are all necessary data fields entered, and in proper format?	Yes	None
Are at least 4 data points present for statistical analysis?	Yes	None
Are at least 8 data points present for an attainment evaluation?	Yes	None
Are detection limits for nondetects ≤ maximum detected value?	Yes	None
Are all data within chart axis limits?	Yes	None

Select Axis Values and Press "Reset Axes" after entering data.

Axis Values			
Time		Concentration	
Min	Max	Min	Max
8	19	Auto	Auto

Reset Axes

Pressing the "Check for Outliers" button to the right will open a worksheet that shows the results of a Dixons's test for outliers.

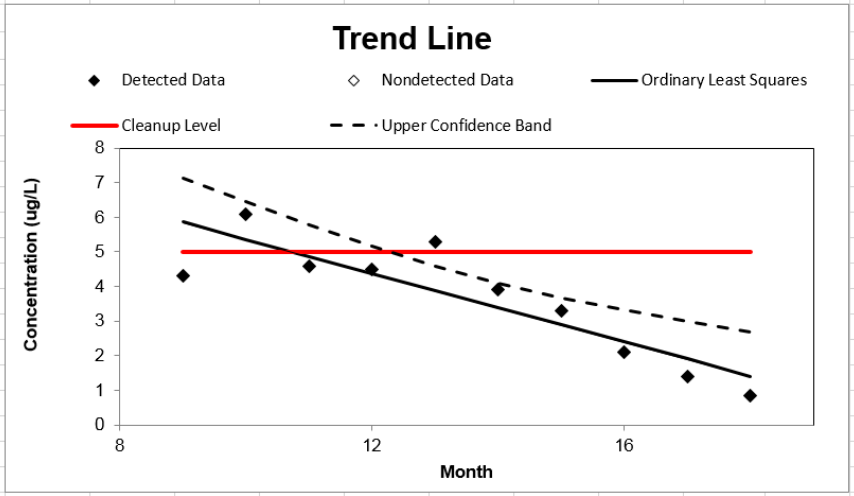
Next Step: Check for Outliers

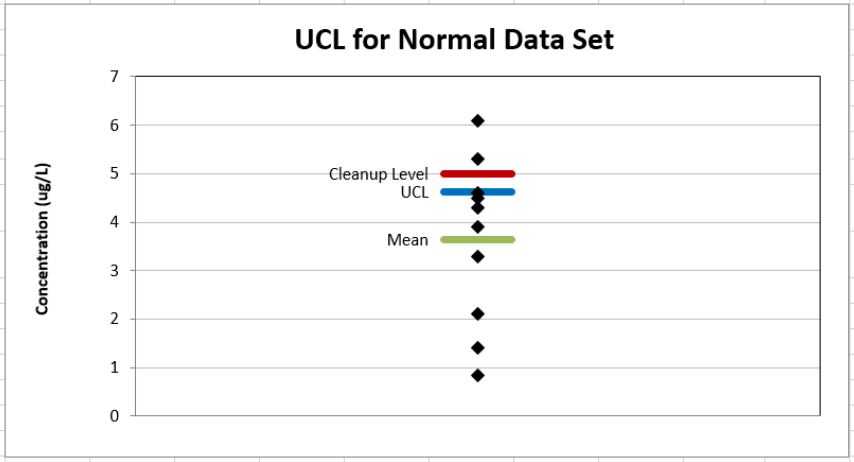
Figure 12. Site & Summary Statistics Screen for Trichloroethene Attainment Example 2, All Ten Data Points Used

Groundwater Statistics Tool

Site & Summary Statistics for Normal Data Sets with Normal Residuals

General Information	
Analyst	DW
Date of Evaluation	4/9/2018
Site Name	Test
Operable Unit	Test
Type of Evaluation	Attainment
Well Name/Number	1
Chemical of Concern	TCE
Concentration Units	ug/L
Cleanup Level	5
Source of Cleanup Level	MCL
Confidence Level	95%
Risk of False Outlier Rejection	1%
Number of Results	10
Outliers present?	No
Number of Non-Detects	0
Trend Analysis	
Trend Type	Normal
Method	Ordinary Least Squares
Is the Upper Confidence Band above the cleanup level?	**NA**
Slope	-0.494
Intercept	10.3
R^2	0.77
Test Result	**Decreasing**
When concentration is predicted to achieve the cleanup level	**NA**
When concentration is predicted to exceed the cleanup level	**NA**
UCL Analysis	
Distribution Type	Normal
Test	Student's t UCL
Mean	3.64
95% UCL	**4.63**
Is the 95% UCL greater than the cleanup level?	**No**





Previous Step: Trend Screen | Previous Step: UCL Screen | Restart: Data Input Screen

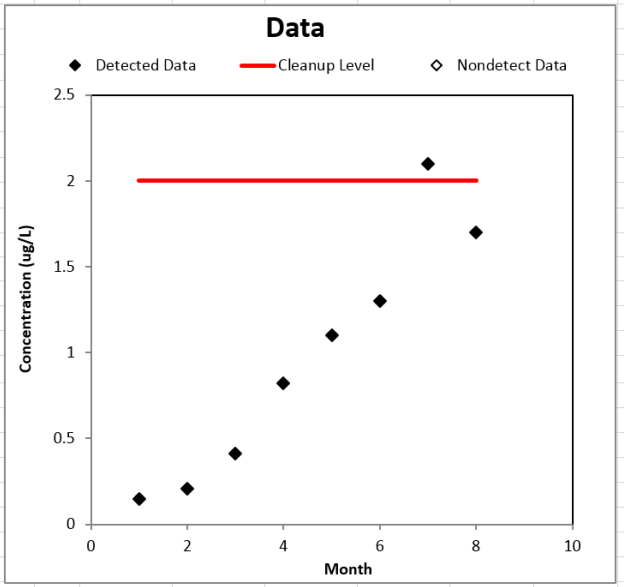
Figure 13. Data Input Screen for Vinyl Chloride Remediation Example 3

Groundwater Statistics Tool

Data input worksheet

Site Name	Test
Operable Unit (OU)	Test
Type of Evaluation	Remediation
Date of Evaluation	4/9/2018
Person performing analysis	DW
Chemical of Concern	VC
Well Name/Number	1
Date Units	Month
Concentration Units	ug/L
Confidence Level Desired	95%
Cleanup Level	2
Source of cleanup level (e.g. MCL or risk-based concentration)	MCL
Risk of False Outlier Rejection	1%
Random Seed (may be left blank)	62311.42188
Significant figures to use	3
Number of data points:	8
Number of detected results:	8
Number of nondetect results:	0
Detection frequency:	100%

Date (Month)	VC Conc. (ug/L)	Data Qualifier	Detected? (Yes or No)
1	0.15		Yes
2	0.21		Yes
3	0.41		Yes
4	0.82		Yes
5	1.1		Yes
6	1.3		Yes
7	2.1		Yes
8	1.7		Yes



Data Review		Recommendations
Are all necessary data fields entered, and in proper format?	Yes	None
Are at least 4 data points present for statistical analysis?	Yes	None
Are at least 8 data points present for an attainment evaluation?	Yes	None
Are detection limits for nondetects ≤ maximum detected value?	Yes	None
Are all data within chart axis limits?	Yes	None

Select Axis Values and Press "Reset Axes" after entering data.

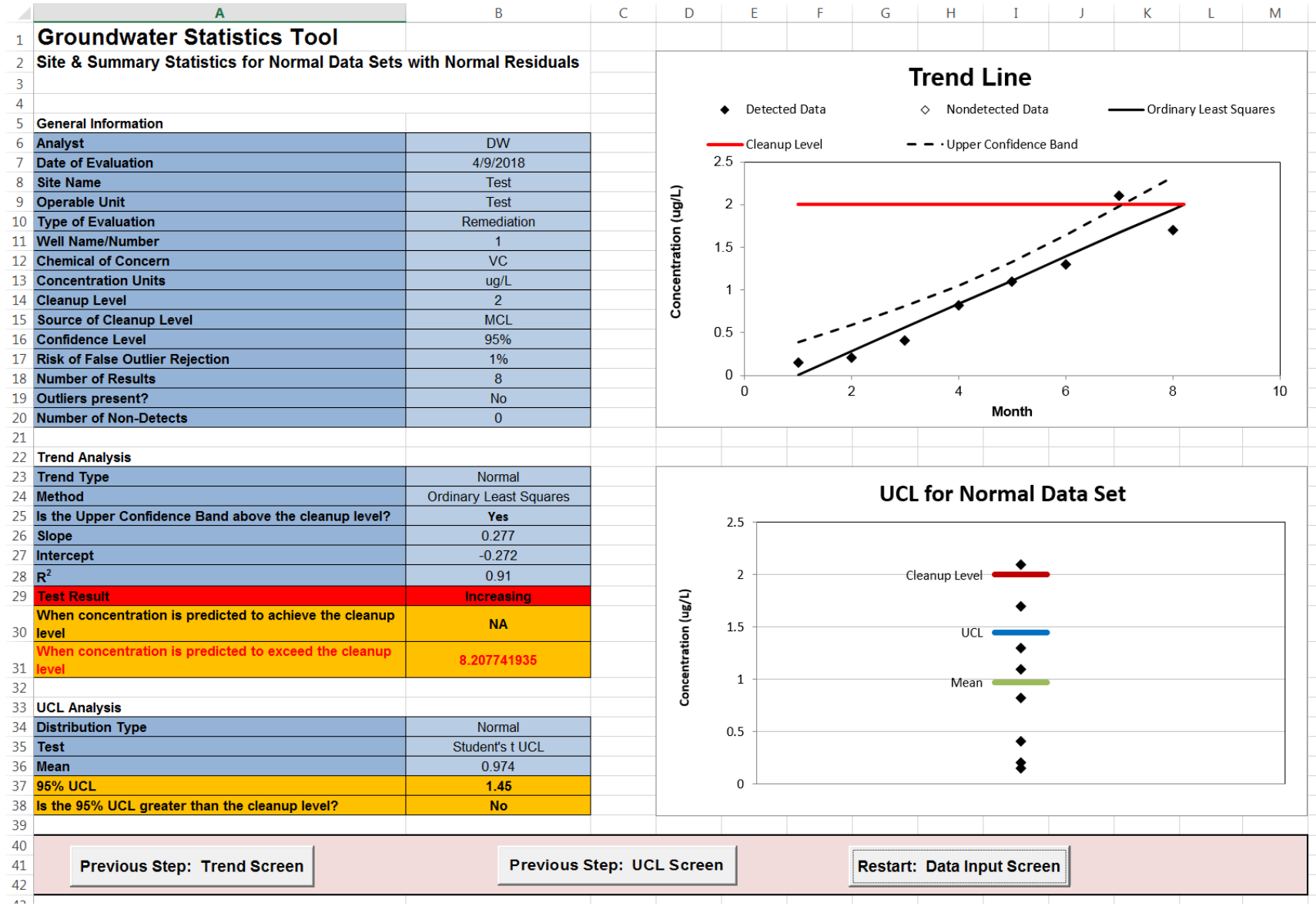
Axis Values			
Time		Concentration	
Min	Max	Min	Max
Auto	Auto	Auto	Auto

Reset Axes

Pressing the "Check for Outliers" button to the right will open a worksheet that shows the results of a Dixons's test for outliers.

Next Step: Check for Outliers

Figure 14. Site & Summary Statistics Screen for Vinyl Chloride Remediation Example 3



Groundwater Statistics Tool

Site & Summary Statistics for Normal Data Sets with Normal Residuals

General Information	
Analyst	DW
Date of Evaluation	4/9/2018
Site Name	Test
Operable Unit	Test
Type of Evaluation	Remediation
Well Name/Number	1
Chemical of Concern	VC
Concentration Units	ug/L
Cleanup Level	2
Source of Cleanup Level	MCL
Confidence Level	95%
Risk of False Outlier Rejection	1%
Number of Results	8
Outliers present?	No
Number of Non-Detects	0
Trend Analysis	
Trend Type	Normal
Method	Ordinary Least Squares
Is the Upper Confidence Band above the cleanup level?	**Yes**
Slope	0.277
Intercept	-0.272
R^2	0.91
Test Result	**Increasing**
When concentration is predicted to achieve the cleanup level	**NA**
When concentration is predicted to exceed the cleanup level	**8.207741935**
UCL Analysis	
Distribution Type	Normal
Test	Student's t UCL
Mean	0.974
95% UCL	**1.45**
Is the 95% UCL greater than the cleanup level?	**No**

Previous Step: Trend Screen | Previous Step: UCL Screen | Restart: Data Input Screen