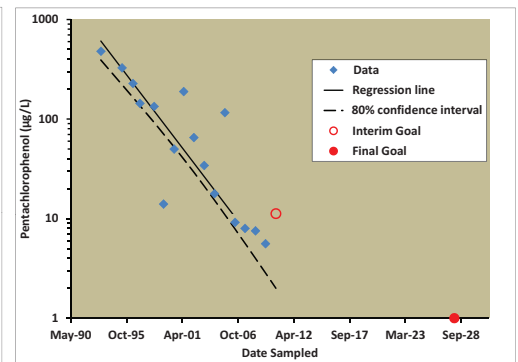
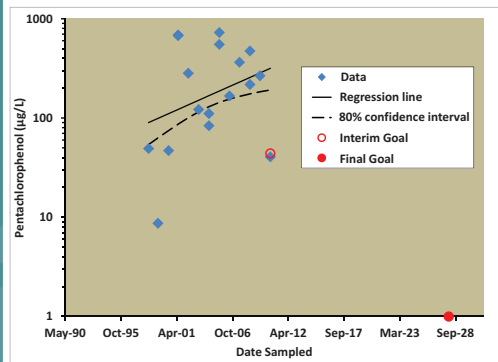


An Approach for Evaluating the Progress of Natural Attenuation in Groundwater



An Approach for Evaluating the Progress of Natural Attenuation in Groundwater

John T. Wilson

*U.S. EPA/NRMRL/ORD/Ground Water and
Ecosystems Restoration Division*

Notice

The U.S. Environmental Protection Agency through its Office of Research and Development conducted the research described here as an in-house effort in collaboration with the Ground Water Forum. The Forum is a group of groundwater scientists that support the Superfund and RCRA programs in each of the ten EPA Regional Offices. This Report has been subjected to the Agency's peer and administrative review and has been approved for publication as an EPA document.

Nothing in this Report changes Agency policy regarding remedial selection criteria, remedial expectations, or the selection and implementation of MNA. This document does not supersede any previous guidance and is intended for use in conjunction with the OSWER Directive 9200.4-17P, *Use of Monitored Natural Attenuation at Superfund, RCRA Corrective Action, and Underground Storage Tank Sites* (U.S. EPA, 1999) and *Performance Monitoring of MNA Remedies for VOCs in Ground Water*. EPA/600/R-04/027 (Pope et al., 2004).

Foreword

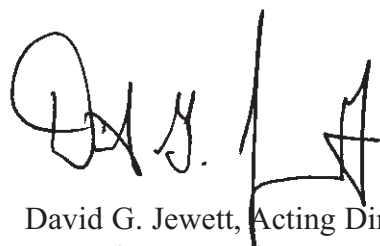
The U.S. Environmental Protection Agency (EPA) is charged by Congress with protecting the Nation's land, air, and water resources. Under a mandate of national environmental laws, the Agency strives to formulate and implement actions leading to a compatible balance between human activities and the ability of natural systems to support and nurture life. To meet this mandate, EPA's research program is providing data and technical support for solving environmental problems today and building a science knowledge base necessary to manage our ecological resources wisely, understand how pollutants affect our health, and prevent or reduce environmental risks in the future.

The National Risk Management Research Laboratory (NRMRL) is the Agency's center for investigation of technological and management approaches for preventing and reducing risks from pollution that threatens human health and the environment. The focus of the Laboratory's research program is on methods and their cost-effectiveness for prevention and control of pollution to air, land, water, and subsurface resources; protection of water quality in public water systems; remediation of contaminated sites, sediments and ground water; prevention and control of indoor air pollution; and restoration of ecosystems. NRMRL collaborates with both public and private sector partners to foster technologies that reduce the cost of compliance and to anticipate emerging problems. NRMRL's research provides solutions to environmental problems by: developing and promoting technologies that protect and improve the environment; advancing scientific and engineering information to support regulatory and policy decisions; and providing the technical support and information transfer to ensure implementation of environmental regulations and strategies at the national, state, and community levels.

Monitored Natural Attenuation (MNA) is widely applied to ground water contamination at hazardous waste sites. Under the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA), MNA is considered to be a remedy like any other remedy. When MNA has been selected as a remedy, concentrations of contaminants in the groundwater are expected to achieve a clean-up goal at the site within a reasonable time frame. At many CERCLA sites, the time by which the goals are to be obtained is specified in the Record of Decision (ROD). At CERCLA sites, the performance of the remedy or combination of remedies that were selected under the ROD is reviewed every five years.

At present, there is no generally accepted approach to evaluate long term monitoring data and establish a time by which clean up goals should be attained. At present, there is no generally accepted approach that can be used to determine whether the extent of attenuation within a particular five year review period is adequate to allow the site to attain the clean-up goal at the site within a specified time frame.

This report presents a simple, statistically based approach for evaluating the progress of natural attenuation from the data collected during site characterization and long term monitoring. The report provides an approach to establish the time that should be required to attain the clean up goals, and an approach to evaluate attenuation within a review cycle to determine whether attenuation within that period of time is adequate to allow the site to attain the ultimate clean-up goal by a specified time.



David G. Jewett, Acting Director
Ground Water and Ecosystems Restoration Division
National Risk Management Research Laboratory

Contents

Notice	ii
Foreword	iii
Acknowledgements	xi
Abstract	xiii
1.0 Introduction	1
1.1. Data analysis best performed before selection of MNA as remedy	4
1.2. Data analysis performed after selection of MNA as remedy	5
2.0 Illustration of the First Phase of Analysis	7
2.1. Identifying the rate law for natural attenuation	9
2.2. Estimating time required to reach clean up goals	10
2.3. Estimating uncertainty in the rate of natural attenuation	11
2.4. Shortcuts to estimate the rate of attenuation	12
3.0 Illustration of the Second Phase of Analysis	13
3.1. Testing whether the reduction in concentration is statistically significant	14
3.2. Testing whether the reduction in concentration is adequate to meet goals	14
3.3. Estimating the probability that the reduction in concentration is not adequate to meet goals ..	15
3.3.1. Establishing a decision criterion	15
3.3.2. Balancing the confidence in the test and the power of the test	16
3.3.3. Applying the decision criterion	17
4.0 Regression as an Alternative Phase Two Analysis	19
4.1. Establishing a decision criterion for regression	19
4.2. Regression with only one sample each year	21
5.0 Putting the Statistical Analysis into a Geohydrological Framework	23
5.1. Monitoring to document trends in attenuation in three dimensional space	23
5.2. Evaluation of whether attenuation is adequate at a site	24
5.2.1. Evaluation of the site as a whole	24
5.2.2. Evaluation of individual wells	27
5.2.3. Uncertainty in estimates of time to attain cleanup goal	28
5.2.4. Interpretation of projections	28
5.2.5. Transformation products from natural attenuation may be a special case	29
5.2.6. Rates of attenuation reported in the literature	30
5.2.7. Dealing with data quality issues	31
6.0 Suggestions and Recommendations	33
7.0 References	35
Appendix A First Phase Analysis, Using Linear Regression to Extract Rate Constants	37
A.1. Use of the spreadsheet to calculate a linear regression	37
A.1.1. Express the sample dates as decimal years	37
A.1.2. Calculate natural logarithms	38
A.1.3. Run the regression	39
A.1.4. Examine the results	40
A.2. Statistical background on linear regression	41
A.2.1. Linear relationship between variables	42

A.2.2. Uniform variance in data	42
A.2.3. Normal distribution of residuals	42
A.3. Using Excel to generate a Q-Q plot and using the One-Sample Kolmogorov-Smirnov Goodness-of-Fit Test	44
A.4. Using ProUCL for a Q-Q plot and Goodness-of-Fit testing.	48
References	49
Appendix B Second Phase Analysis	53
B.1. Statistical Approach	53
B.1.1. Statistical background: Use of the Student's t -statistic	53
B.1.2. Theoretical basis of the statistical comparison of means	54
B.1.3. Use of a spreadsheet to calculate the difference of means	55
B.1.4. Modifying the spreadsheet to accommodate samples	61
B.2. Independence of samples and seasonal effects.	62
B.3. Selecting the appropriate value of α	63
Reference.	65

Figures

Figure 1.	Distribution of TCE, <i>cis</i> -DCE and Vinyl Chloride in the most contaminated well in a transect of monitoring wells arranged between a source of contamination of TCE in ground water and Lake Michigan.	7
Figure 2.	Monitoring record for concentrations of TCE in well MW-3B, a Point of Compliance Well in a transect of monitoring wells.	9
Figure 3.	Comparison of the means for the first year of the review cycle (2001) and the final year of a five year review cycle (2006) to the clean up goals.	13
Figure 4.	A comparison of the concentrations in samples in the initial and final year of the review cycle to the interim goals in the final year of the cycle (2006) to be adequate for the long term clean up goal (5 µg/L in 2017).	16
Figure 5.	A comparison of the regression line through the natural logarithm of the concentrations of TCE, and the 80% confidence interval on the line to interim goal at the end of the review cycle (2007) and the long term clean up goal (5 µg/L in 2017).	20
Figure 6.	Effect of number of samples on the confidence belts on regression lines fit through the data.	21
Figure 7.	Projected 60% confidence belts on the regression of the natural logarithm of concentration of vinyl chloride in MW-3A on date.	28
Figure A.1.	A spreadsheet to calculate the decimal date of sampling and the natural logarithm of the concentration of TCE.	38
Figure A.2.	The Regression dropdown menu under the Data Analysis tab in the Tools menu.	39
Figure A.3.	The input menu for linear regression.	40
Figure A.4.	The summary output of the linear regression.	41
Figure A.5.	Fit of a linear trend and an exponential trend to concentration of TCE.	43
Figure A.6.	Fit of a linear trend to the natural logarithm of concentration of TCE.	44
Figure A.7.	Report in Excel 2003 of the Residuals and Standard Residuals from the Linear Regression of the example data set.	45
Figure A.8.	Calculation of the values of the standard normal distribution (z-distribution) corresponding to the rank and quantile of the standard residuals.	46
Figure A.9.	Normal Probability Plot (also called a Q-Q plot) comparing the distribution of the residuals from the regression to the normal probability distribution.	47
Figure A.10.	Data entry form in an application to perform the Kolmogorov-Smirnov Test.	48
Figure A.11.	Results returned from the Kolmogorov-Smirnov test.	49
Figure A.12.	Using ProUCL to generate a Q-Q plot.	50
Figure A.13.	Using ProUCL to evaluate goodness-of-fit to a normal distribution.	51
Figure B.1.	Populating the spreadsheet <i>Evaluation of MNA</i> to calculate the mean concentrations of contaminant in the first and in the final year of the review cycle.	55
Figure B.2.	Calculating the standard deviation of the means, the difference between the means, and the standard deviation of the difference between the means.	56

Figure B.3. Calculating the difference between the means necessary to be statistically significant at a predetermined probability of error. 57

Figure B.4. Setting interim clean-up goals for the final year of the review cycle. 59

Figure B.5. Calculating the mean of the interim goals, the standard deviation of the interim goals, the difference between the mean of samples and the mean of the interim goals, and the standard deviation of the difference between the mean of samples in the final year and the mean of the interim goals. 60

Figure B.6. Evaluation whether attenuation is adequate to attain goals. 61

Figure B.7. Data input screen and subscreen for G*Power 3.1.2. 64

Tables

Table 1.	Maximum concentrations of TCE, <i>cis</i> -DCE and Vinyl Chloride in a transect of monitoring wells (12/11/2000 to 10/03/2003).	8
Table 2.	Second phase of analysis.	26
Table 3.	Using the first phase of analysis to project when concentrations of TCE, <i>cis</i> -DCE and Vinyl Chloride in monitoring wells will attain the clean up goal.	27
Table 4.	Typical rates of attenuation of concentrations over time in monitoring wells.	31

Acknowledgments

The Groundwater Forum is a group of ground-water scientists that support the Superfund and RCRA programs in each of the ten EPA Regional Offices. This report was developed for and in close consultation with the Groundwater Forum. Luanne Vanderpool with U.S. EPA Region 5 managed the review of the report for the Groundwater Forum, and provided the EPA internal peer review. External peer reviews were provided by Thomas McHugh and Charles Newell of Groundwater Services Inc, Houston, Texas; Yue Rong of the California Regional Water Quality Control Board, Los Angeles, California; and Aristeo (Resty) Pelayo of the Wisconsin Department of Natural Resources, Madison, Wisconsin.

Abstract

Monitored Natural Attenuation (MNA) remedies monitor the results of natural processes and utilize these data to predict future conditions in the aquifer. Performance monitoring to evaluate MNA effectiveness is a critical element for MNA remedies.

Monitoring over time insures that the future behavior of the plume is consistent with past behavior and that the risk of exposure to the contaminants is managed. The trend of contaminant concentrations over time in a particular well can be used to forecast the future concentrations in that well and predict when concentrations will attain a selected concentration level. The purpose of this document is to present a simple, statistically based approach for evaluating the progress of natural attenuation from the data collected during site characterization and long term monitoring. The intended audience is technical professionals that actually perform the data analyses (i.e., hydrogeologists, engineers) as well as project managers who review those analyses and/or make decisions based on those analyses.

Monitored Natural Attenuation (MNA) was a component of more than 20 percent of remedies implemented between 1982 and 2005 at National Priority List (NPL) sites where groundwater is contaminated (U.S. EPA, 2007). In the interval 2005 through 2008, MNA was a component of more than 18% of NPL sites where groundwater is contaminated (U.S. EPA, 2010). The percentage of sites using MNA is likely to increase as source areas are depleted by more aggressive remedies and MNA is used to remediate remaining contamination.

MNA remedies monitor the results of natural processes and utilize these data to predict future conditions in the aquifer. MNA remedies for contaminated groundwater are expected to demonstrate that MNA will “... *attain cleanup levels (or other remedial action objectives) in a time frame that is reasonable when compared to the cleanup time frames of the other alternatives and when compared to the time frame of the anticipated resource use*” (U.S. EPA, 1999, p. 9-7). This statement defines key expectations which EPA has for MNA remedies and is a measure of success for all MNA remedies. This paper describes methods which can be used to define “... *a timeframe that is reasonable ...*” and methods which quantify demonstrations that MNA will achieve the remedial action objectives.

Performance monitoring to evaluate MNA effectiveness is a critical element for MNA remedies. One objective for the performance monitoring program of an MNA remedy is to “Demonstrate that natural attenuation is occurring according to expectations” (U.S. EPA, 1999, p. 22). The expectations generally include being able to achieve the concentration based clean-up goals for the contaminants in the groundwater at the site. A Record of Decision (ROD) for a site regulated under

the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) may specify a particular date by which specific clean-up goals must be reached. Whether a specific date has been specified or whether the cleanup time must simply be reasonable, the approach presented in this document describes a methodology using simple statistics which objectively demonstrates whether a MNA remedy is proceeding at a rate which will achieve the goals specified in the ROD and is meeting expectations. The method is specifically intended to support the five-year remedy review process under CERCLA; however, it is applicable to any site that has concentration-based clean up goals for the contaminants, and a date (whether specified or simply “reasonable”) by which the goals must be attained.

Monitored Natural Attenuation (MNA) is generally considered in enforcement actions regulated under CERCLA to be a remedy like any other remedy. The Record of Decision (ROD) will specify clean up goals for each contaminant at the site. These goals are usually expressed in terms of the maximum concentration of the contaminants in soil or groundwater that will be tolerated at the site when the remedy is complete.

Monitoring over time insures that the future behavior of the plume is consistent with past behavior and that the risk of exposure to the contaminants is managed. The trend of contaminant concentrations over time in a particular well can be used to forecast the future concentrations in that well and predict when concentrations will attain a selected concentration level. The approach presented in this paper can be used to estimate cleanup times by MNA even where clean up goals are not specified by time or concentration.

The approach is intended to be used in the context of existing U.S. EPA guidance and recommendations for MNA. These U.S. EPA documents can be found at <http://www.epa.gov/oust/cat/mna.html> and an associated link at <http://www.epa.gov/ada/gw/mna.html>. Two of the U.S. EPA documents are widely applicable: *Use of Monitored Natural Attenuation at Superfund, RCRA Corrective Action, and Underground Storage Tank Sites* (U.S. EPA, 1999) and *Performance Monitoring of MNA Remedies for VOCs in Ground Water* (Pope et al., 2004). U.S. EPA (1999) lays out a general framework for the appropriate implementation of MNA at EPA enforcement actions. Pope et al. (2004) summarize that comprehensive framework in the following steps:

- 1) demonstrate that natural attenuation is occurring according to expectations,
- 2) detect changes in environmental conditions (e.g., hydrogeologic, geochemical, microbiological, or other changes) that may reduce the efficacy of any of the natural attenuation processes,
- 3) identify any potentially toxic and/or mobile transformation products,
- 4) verify that the plume(s) is not expanding downgradient, laterally or vertically,
- 5) verify no unacceptable impact to downgradient receptors,
- 6) detect new releases of contaminants to the environment that could impact the effectiveness of the natural attenuation remedy,
- 7) demonstrate the efficacy of institutional controls that were put in place to protect potential receptors, and
- 8) verify attainment of remediation objectives.

Performance Monitoring of MNA Remedies (Pope et al., 2004) briefly discusses the use of statistics to evaluate trends. In their Section 3.3.34 they state:

Statistical methods are also available to facilitate analysis and comparison of trends by considering data variability through time. For instance, changes in contaminant concentrations over space or time can be used to calculate attenuation rates, and the variability associated with those rates can be quantified with confidence intervals about the rates. These confidence intervals can be used to determine the likelihood of attaining site-related remedial goals. If all values of the attenuation rate falling within the confidence intervals lead to predictions that site remedial goals will be attained in the desired time frame, then confidence that MNA can attain remedial goals is increased.

At many sites, the evaluation of trends in concentrations stops with an evaluation of whether natural attenuation is occurring, and the evaluation does not proceed to the next step of evaluating whether MNA meets expectations. This Report provides an approach to carry out the first step in the comprehensive evaluation of the performance of MNA, which is to demonstrate that natural attenuation is occurring according to expectations.

In CERCLA, the clean-up goals must be met throughout the plume of contamination. Progress should not be evaluated at only a few selected wells. Care must be taken to avoid making generalizations about clean-up progress without careful consideration of the entire monitoring network. While the example used to demonstrate the methodologies outlined in this paper focuses on a single well, the methodologies should be applied to the entire monitoring network. Many sites will have problematic wells where the trends in concentration within the five year review cycle are not adequate. Depending on circumstances, the trends in the problematic wells may or may not indicate that MNA is inadequate for the plume as whole. Interpretation of data from problematic wells is discussed in Section 5.2.

The approach presented in this document is divided into two phases. The first phase is

best undertaken before MNA is selected as part of the remedy at a site. The second phase is undertaken as a part of performance monitoring after the selection of MNA as a remedy.

In the first phase, the trend in concentrations at individual wells is evaluated to estimate the date when concentrations can be expected to meet the clean up goal at each well. The approach in the first phase can be used to evaluate in a general manner whether natural attenuation processes have a reasonable likelihood of reaching concentration based goals in a reasonable time period. If a date to attain the clean up goals has not been established at a site, the first phase also provides one approach that can be used to select a date (see Sections 1.1 and 5.2.3).

In the second phase of analysis, data are evaluated over the interval of time in a review cycle, such as the five-year review cycle that is common in CERCLA. The attenuation in concentration between the initial and final years of the interval is compared to the attenuation that would be required in order to attain the clean up goal by the specified date. If the concentration in the final year of the review cycle is higher than the concentration that would be required to attain the goal by the specified date, the second phase estimates the probability that natural attenuation will not meet the clean up goal.

Before this approach (or any other approach to evaluate MNA) is applied to data generated from monitoring wells, the user should consider the representativeness of the existing wells in portraying the groundwater conditions, and whether the existing monitoring wells adequately characterize the plume vertically and horizontally. This type of data analysis and forecasting is only as reliable as the representativeness of data that is subjected to the analysis. Subtle deterioration in the condition of the well, incorrect interpretations of groundwater flow directions or changes in the groundwater direction over time can be falsely interpreted as “attenuation”. The approach presented in this document is only appropriate when it has been

determined that the monitoring well network accurately represents the ground water flow system. Consult *Performance Monitoring of MNA Remedies* (Pope et al., 2004) for recommendations on the design of a monitoring system.

The purpose of this document is to present a simple, statistically based approach for evaluating the progress of natural attenuation from the data collected during site characterization and long term monitoring. The intended audience is technical professionals that actually perform the data analyses (i.e., hydrogeologists, engineers) as well as project managers who review those analyses and/or make decisions based on those analyses. In this document the two phases of the approach are described and illustrated with an example. Since the approach outlined in this document uses statistical methods, even the simplified description in the body of the report is somewhat technical. Non-technical readers may be best advised to read Sections 1 and 5 first to understand the kind of decisions to be made using the methods described in this paper. The two appendices are more technical and provide additional information regarding the statistical methods used in this approach as well as detailed step by step instructions to apply the methods to a data set.

Section 1 introduces the approach for evaluating the progress of Natural Attenuation.

Section 2 introduces the statistics which are used in Phase 1 to perform the evaluation of the progress of MNA and describes the application of these methods to characterize the behavior of contaminants at the site before MNA is selected as a remedy for the site. Methods are presented to evaluate trends in concentrations over time and to estimate the time required to reach clean-up goals. The statistical methods are described in much more detail in the appendixes which accompany this report, but Section 2 is a statistical discussion and is somewhat technical.

Section 3 applies the statistical methods used in Phase 2 to evaluate data from a long-term

monitoring program in order to determine whether the progress of natural attenuation is sufficient to achieve the clean-up goals set for the site. The methods are built on the *t*-test for the difference of means. The statistical methods are described in much more detail in the appendixes which accompany this report, but Section 3 is a statistical discussion and like Section 2 is somewhat technical.

Section 4 describes the use of linear regression techniques to evaluate the progress of Natural Attenuation and for documenting the level of confidence which can be placed in decisions made from the data. This is an alternative approach to the method using the *t*-statistic presented in Section 3.

Section 5 describes the application of the decision process for evaluating the progress of Natural Attenuation using the results from methods described in Sections 2 through 4. Section 5 uses technical terms from earlier sections, but presents more general discussion regarding the performance of MNA at the site. These are issues which decision makers must recognize and consider even if they do not perform the statistical analyses themselves.

1.1. Data analysis best performed before selection of MNA as remedy

The first phase of data analysis determines, based on the available monitoring record, whether natural attenuation processes appear to be capable of reducing the concentrations of contaminants to the clean up goals within a reasonable time period. This evaluation requires sufficient monitoring data to indentify temporal trends in contaminant concentration. This first phase of evaluation should be conducted before MNA is selected as part of the remedy at a site.

The first phase of data analysis can be used to meet the particular requirements set out in the OSWER Directive (U.S. EPA 1999) regarding the use of MNA. Text quoted from the Directive is in bold font. The Directive outlined **a three-tiered approach to site characterization**. The first line of evidence presented in

the Directive was: **(1) Historical groundwater and/or soil chemistry data that demonstrate a clear and meaningful trend of decreasing contaminant mass and/or concentration over time at appropriate monitoring or sampling points**. In addition, the Directive states that ... **EPA expects that documenting the level of confidence on attenuation rates will provide more technically defensible predictions of remedial timeframes and form the basis for more effective performance monitoring programs**. The Directive also states: **Statistical confidence intervals should be estimated for calculated attenuation rate constants (including those based on methods such as historical trend data analysis ...)**.

The first phase of the approach presented in this document uses regression analysis to evaluate natural attenuation. Regression analysis provides rate constants for attenuation that provide a precise definition of a “clear and meaningful trend of decreasing contaminant concentration over time.” Because higher rate constants mean faster remediation time frames, simple calculations in a spread sheet are used to determine whether the rate constants are different from zero at some predetermined level of statistical confidence and to extract the statistical confidence interval on the rate of attenuation.

At many sites where MNA has been selected as remedy, a concentration-based goal for cleanup has been identified, but the date to attain that goal is not identified. Ideally, the target date to attain the goal should be established based on factors such as the timeframe for anticipated future use of the resource, a comparison to the time that would be required for alternative remedial techniques to attain the goal, and the reliability of existing institutional controls and other controls on exposure to contamination.

One option to establish a target date is to use regression analysis to extrapolate natural attenuation in the historical record and predict the time when the goal can be expected to be attained. This date can then be compared to the requirements imposed by site specific factors

to determine if it is appropriate. This option is straightforward, but it has a serious implication. If the rate of attenuation continues into the future at the same rate as in the past, on average, half the time the true date to attain the goal will be further in the future than the target date, and half the time continued long term monitoring will reveal that MNA is not adequate to attain the final goal by the target date. If regression analysis is used to select a target date, that date should acknowledge and make provision for the uncertainty in the regression (see Section 5.2.3).

1.2. Data analysis performed after selection of MNA as remedy

The second phase would apply to evaluations after approval of MNA as a remedial measure. Fitting one trend to all the data assumes that the rate of natural attenuation does not change over the history of the site. However, the rate of attenuation can change for a variety of reasons. The monitoring record is subject to temporal variations in concentrations that create temporary or “apparent” trends. High water tables can inundate residual contamination that was previously in the unsaturated zone, producing slugs of contaminants at high concentration. At some sites, critical substrates are required by the bacteria that degrade the contaminants. Examples would include molecular hydrogen and acetate used by *Dehalococcoides* bacteria to degrade trichloroethylene, or sulfate used by bacteria to degrade benzene. These substrates can be depleted over time, which may change the rate of biodegradation. These interactions and more are discussed in greater detail in *Performance Monitoring of MNA Remedies* (Pope et al., 2004). Statistical analysis is one approach to recognize a change in the trend in concentrations over time.

Active source control or source remediation will reduce concentrations of contaminants, and the benefit of this active treatment will be included in the overall trend in contaminant concentrations. Biostimulation with growth substrates and nutrients or bioaugmentation with active

microorganisms will increase the rate of biodegradation. Some vendors even talk about enhanced natural attenuation. These active measures will interfere with a statistical evaluation of MNA. It is best to restrict an evaluation of natural attenuation to the time periods before active remediation has been implemented, or to time periods after the benefits of the active remediation have been realized.

The purpose of the second phase of analysis is to determine whether MNA continues to perform as expected, and whether the recent performance of MNA at the site is adequate to meet the clean up goals. Depending on the outcome, the analyses in the second phase might support a decision to allow a MNA remedy to proceed as defined in the ROD. Depending on the outcome, the analyses may also identify a need for additional sampling and data evaluation, or identify the need for additional active remediation at a site.

Illustration of the First Phase of Analysis

Figure 1 shows the location of clusters of monitoring wells at a CERCLA enforcement action where MNA has been selected as part of the remedy. Figure 1 also presents the highest concentrations of contaminants observed in the well clusters. The highest concentrations

of contaminants are in the MW-3 well cluster. Table 1 presents the highest concentrations of contaminants observed in the individual wells. The well with the highest concentrations of vinyl chloride (the compound of greatest concern) is well MW-3B, the middle well of the

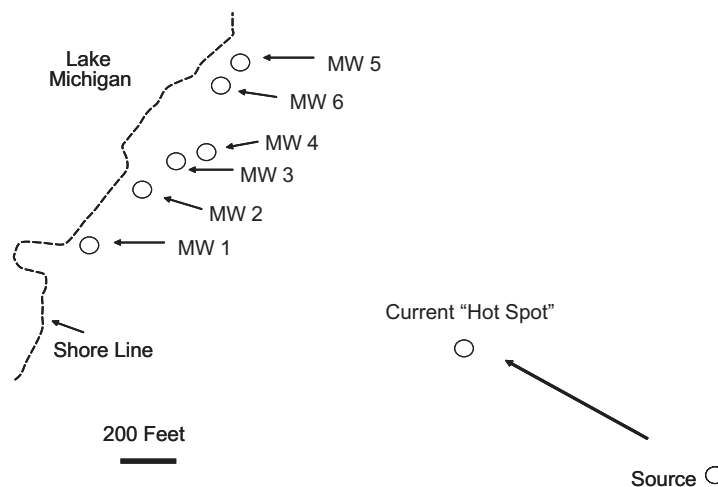
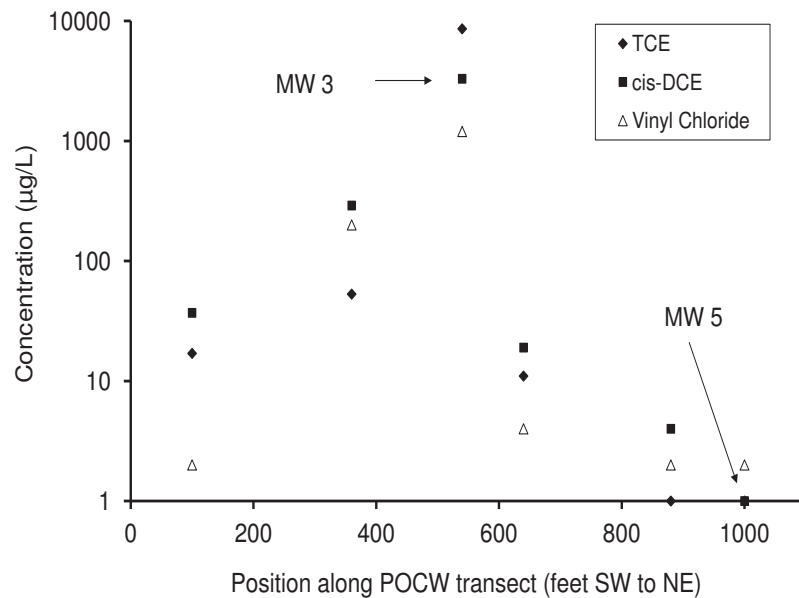


Figure 1. Distribution of TCE, cis-DCE and Vinyl Chloride in the most contaminated well in a transect of monitoring wells arranged between a source of contamination of TCE in ground water and Lake Michigan.

Table 1. Maximum concentrations of TCE, *cis*-DCE and Vinyl Chloride in a transect of monitoring wells (12/11/2000 to 10/03/2003).

Monitoring Well	Depth	TCE µg/L	<i>cis</i> -DCE µg/L	Vinyl Chloride µg/L
MW-1B	Shallow	18	4	2
MW-1A	Deep	17	37	2
MW-2B	Shallow	10	3	2
MW-2A	Deep	53	290	200
MW-3C	Shallow	8600	3300	1200
MW-3B	Intermediate	2300	3200	2200
MW-3A	Deep	20	1900	1500
MW-4C	Shallow	18	12	2
MW-4B	Intermediate	11	19	4
MW-4A	Deep	2	1	4
MW-5B	Shallow	1	1	2
MW-5A	Deep	1	1	2
MW-6		1	4	2

MW-3 well cluster. Figure 2 presents long term monitoring data of TCE concentrations in well MW-3B. This data set will be used to illustrate the first phase of the analysis, determining whether natural attenuation processes appear to be capable of reducing the concentrations of contaminants to the clean up goals within a reasonable time period.

Changes in the direction of ground water flow can cause the centerline of a plume to shift away from a monitoring well, and give the false impression that concentrations in the plume are declining. The procedures in this Section are only appropriate when it has been determined that ground water flow to the monitoring well is stable. For this example, assume that the flow of ground water at the site has been evaluated

and the direction of flow has not changed to an appreciable extent.

Performance Monitoring of MNA Remedies (Pope et al., 2004) emphasizes the need for evaluating the monitoring well network to determine if it is sufficient for MNA. At any site where this statistical approach might be applied, the monitoring network should first be evaluated to determine that it is sufficient to that purpose. For this example, assume that the monitoring network has been evaluated and determined to be sufficient to document the performance of MNA and it is appropriate to move to the next step of evaluating trends in concentrations.

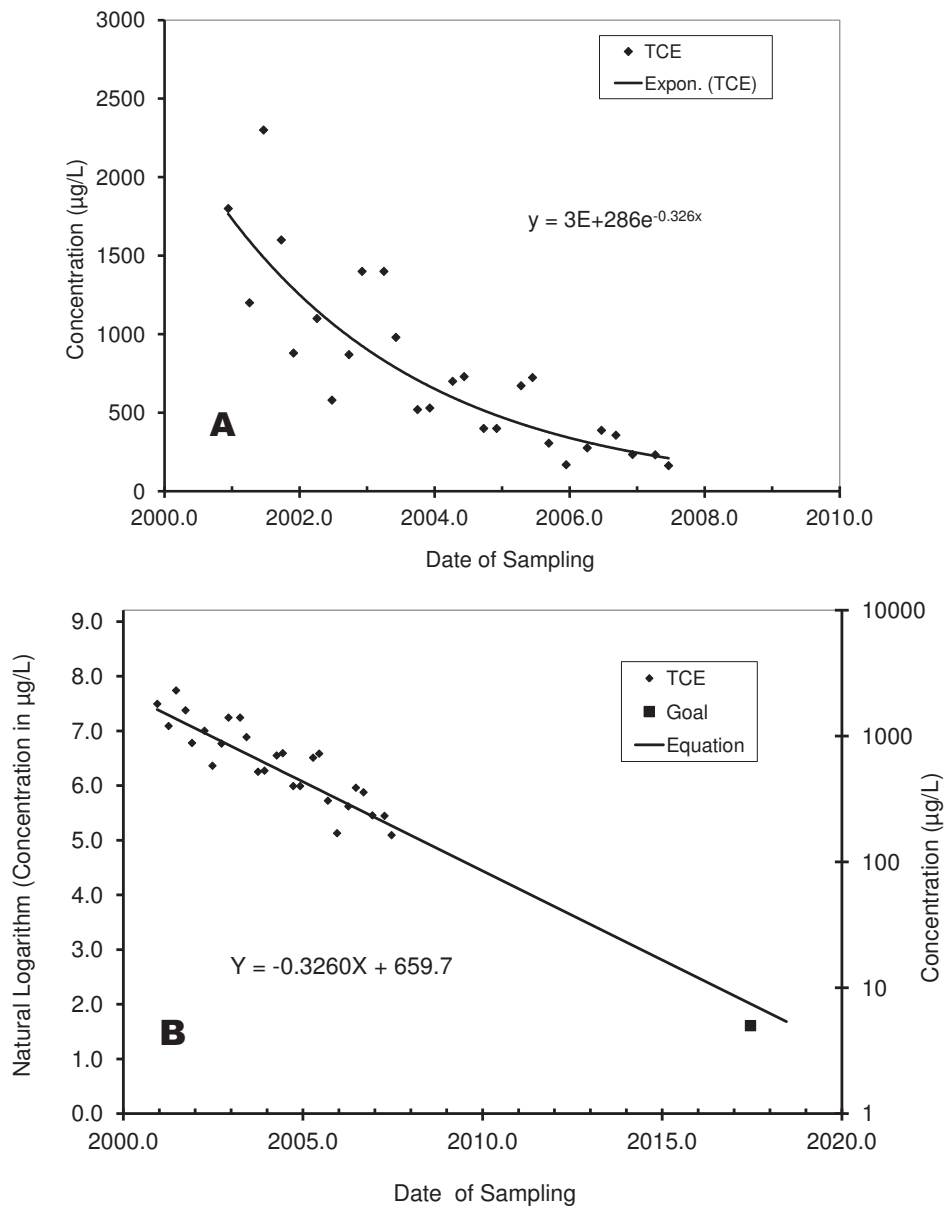


Figure 2. Monitoring record for concentrations of TCE in well MW-3B, a Point of Compliance Well in a transect of monitoring wells.

2.1. Identifying the rate law for natural attenuation

Panel A of Figure 2 presents the monitoring data on an arithmetic scale, where concentration is plotted as a function of time. Over seven years of monitoring, the concentration of TCE (the parent compound) declined from 2,300 µg/L to 163 µg/L. The solid curved line is an exponential trend line that is fit to the data. In this example the fit of the trend line is in reasonable agreement with the data. Notice

that the dates are plotted as decimal years. The slope of the line is the instantaneous rate of attenuation (dC/dt) in µg/L per year. Note that as the concentrations decline, so does the rate of attenuation.

These data follow what is known as first order rate law. The rate of attenuation is described in Equation 1, where r is the rate of change in concentration with time (µg/L per year), C is the concentration of contaminant (µg/L), and

k is the rate constant. The unit for a first order rate constant is reciprocal time. In this case, the unit is (per year).

$$r = dC/dt = k[C]^1 = k[C] \quad 1$$

A familiar example of attenuation under a first order rate law is the decay of radioactive nuclides.

The rate of change in concentration at any instant in time under a first order rate law is proportional to the concentration at that instant ($r = k[C]$). This rate law is called “first” order because the exponent on the concentration C of contaminant is one. The change in the concentration of contaminant in Equation 1 can be integrated over time to produce Equation 2, where t is an interval of time, C is the final concentration at the end of the interval, and C_0 is the initial concentration at the beginning of the time interval.

$$\ln[C] = kt + \ln[C_0] \quad 2$$

If the natural logarithms of the concentrations are plotted as a function of time, the slope of the line is k .

If the exponent on the concentration of contaminant in the rate law were zero, the rate law would be “zero” order and attenuation would follow Equation 3.

$$r = dC/dt = k'[C]^0 = k' \quad 3$$

Under a zero order rate law, the rate of degradation ($\mu\text{g/L}$ per year) is a fixed number, the zero order rate constant k' .

2.2. Estimating time required to reach clean up goals

Panel B of Figure 2 compares the example data on a logarithmic scale. The y-axis on the left side of the figure plots the data as the natural logarithm of the concentrations in $\mu\text{g/L}$. For convenience of comparison, the y-axis on the right side presents the actual concentrations plotted on a logarithmic scale. The data in Panel B appear to fall along a straight line.

The general form for an equation that follows a straight line is presented as Equation 4, where (a) is the slope of the line and (b) is the Y intercept, the value of Y when $X=0$.

$$Y = a * X + b \quad 4$$

If attenuation follows a first order rate law, a plot of the natural logarithm of concentrations on elapsed time will lie along a straight line. The log-transformed data in Panel B of Figure 2 appear to lie along a straight line. A linear regression was used to fit an equation that described the relationship between the natural logarithm of the concentrations of TCE and the date of sampling. The slope of the regression line is the first order rate constant for the change in concentration over time. Details of fitting the line are discussed in Appendix A.

The equation for the line as presented in Panel B of Figure 2 is Equation 5 below, where Y is the natural logarithm of the concentration of TCE in $\mu\text{g/L}$ and X is the date in decimal years.

$$Y = -0.3260 * X + 659.7 \quad 5$$

Solving Equation 4 for X provides Equation 6.

$$X = (Y - b) / (a) \quad 6$$

Solving Equation 5 for X provides Equation 7.

$$X = (Y - 659.7) / (-0.3260) \quad 7$$

For purposes of illustration, we will assume that the desired clean up goal for TCE is $5 \mu\text{g/L}$, and we will assume that it is desired that this goal will be attained by 2017. The natural logarithm of 5 is 1.609. Substituting 1.609 for Y in Equation 7 predicts that natural attenuation will attain the goal by the year 2018.7, which is only slightly later than the desired cleanup date of 2017. The same relationships are presented graphically in Panel B of Figure 2.

As described in Equation 7, the date by which the concentration based cleanup goal must be attained is “key” to the evaluation of the progress of MNA. The statistical evaluation of the progress of MNA in a review cycle is

determined by only four parameters: the date when cleanup is to be attained, the cleanup goal, the historical concentrations of the contaminant and the rate of attenuation.

Note that the value of Y in Equation 7 when X is zero is 659.7 (see the equation in Panel B of Figure 2) which corresponds to a TCE concentration at year zero of 3×10^{286} $\mu\text{g/L}$ (Panel A of Figure 2). This impossibly high value for the initial concentration of TCE was calculated from the regression because the zero value for the Date of Sampling was the beginning of the Common Era. If the regression had been performed on the date expressed in years since the date of the first sample, the value of Y in Equation 7 would have been near to the concentration in the first samples taken.

Trends in contaminant concentrations over time have been reviewed at hundreds of hazardous waste sites. A first order rate law is almost always a better fit than a zero order rate law. The calculations and procedures in this document will assume that attenuation at the site of interest follows a first order law. Equations 4 and 6 can be used to compare the long term performance of natural attenuation to the established clean up goals, or they can be used to predict a time by which the goal may be attained.

No real data set is ever a perfect match to the theoretical assumptions for a particular statistical test, and a certain amount of judgment is required to select the most appropriate data and the most appropriate statistical procedure. In the example data in Panel B of Figure 2, the apparent good agreement with a straight line indicates that (1) the rate of attenuation in concentration follows a first order rate law and (2) the first order rate constant did not change over the time period in the monitoring record. Both these conditions must be true if the calculations in this Section are to be used to forecast a time in the future at which the concentration will attain the clean up goal.

In the illustration, the entire record for the monitoring wells was evaluated. The entire record should be evaluated unless there is a significant change in conditions at the site that justifies using some smaller portion of the record. If the time period in the record includes efforts at source removal or source remediation, the argument could be made that the best data to evaluate the progress of natural attenuation is the portion of the record that occurs after active source control was completed. If the source was remediated by a pump and treat system, the best data to evaluate natural attenuation would be the portion after pumping ceased and the water table returned to its natural condition. In the illustration, there are seven years of quarterly monitoring data. This amount of monitoring data is not always available when the feasibility of MNA as a remedial action needs to be evaluated. However sufficient data should be available to demonstrate a clear and meaningful trend of decreasing contaminant mass and/or concentration over time as specified in the OSWER Directive (U.S. EPA 1999).

If visual examination indicates that a plot of the natural logarithm of concentrations on time is best fit by a curved line, that observation would indicate that either (1) the rate of attenuation does not follow a first order rate law or (2) the first order rate constant is changing over time. If this is the situation, the calculations in this approach should not be used to forecast a time in the future at which the concentration will attain the clean up goal.

2.3. Estimating uncertainty in the rate of natural attenuation

Appendix A of *Performance Monitoring of MNA Remedies* (Pope et al., 2004) discusses a variety of factors that add uncertainty to an estimate of the rate of natural attenuation. The OSWER Directive (U.S. EPA 1999) specifies that this uncertainty in the rate constant be described as a confidence interval on the rate constant. Many evaluations of trends in MNA will express the uncertainty in the fitted trend line as the Coefficient of Determination (r^2).

This is probably done because r^2 is a familiar statistic, and because it is readily obtained from the computer application used to create the figure showing the trend line. The r^2 statistic is not related in a simple way to the confidence interval on the rate of attenuation, and should not be used to evaluate the uncertainty in the trend line. The spreadsheet used to fit the line through the data also has the capacity to put statistical confidence intervals on the slope of the line. Details for this process are illustrated in Appendix A. For the data set in Panel B of Figure 2, the first order rate constant for attenuation of concentrations of TCE was 0.326 per year. As described in Appendix A (Figure A4 and section A.1.4), the rate constant was no slower than 0.281 per year at 90% confidence.

2.4. Shortcuts to estimate the rate of attenuation

The Panels in Figure 2 were generated as charts in Microsoft Excel 2003. The spreadsheet, like many other spreadsheets, allows the user to conveniently fit trend lines to a data set in a chart, without going through the procedures using a spreadsheet to fit a regression line to data as is illustrated in Appendix A. Note, this shortcut approach does not produce a confidence interval for the rate constant as is done using a spreadsheet to calculate the regression.

Options in the chart menu were used to fit an exponential trend line to the data in Panel A of Figure 2. The resulting equation for the line is $Y = 3E + 286e^{-0.326X}$. The date (X in the equation) is an exponent, and the coefficient of X is -0.326. Similarly, options in the chart menu were used to fit a linear trend line to the natural logarithm of the concentration data as shown in Panel B of Figure 2. As discussed earlier, the equation for the line is $Y = -0.3260X + 659.7$. The date is a factor instead of an exponent, but the coefficient of X is the same. In both equations, the coefficient of the time elapsed is the first order rate constant for change in concentration over time. Because time was plotted in decimal years, the unit for the rate constant is reciprocal years. The negative of the coefficient

is the first order rate constant for attenuation over time. The rate constant is negative because concentrations are decreasing.

3.0

Illustration of the Second Phase of Analysis

The same data will be used to illustrate the second phase of the evaluation. Panel A of Figure 3 identifies the data that are involved in the comparison. The mean of concentrations of TCE determined in water samples acquired in each of the four quarters of 2001 will be compared to the mean of concentrations in each of the four quarters of 2006. The data and the means are enclosed in dashed lines. The mean concentration of TCE in 2001 was 1404 $\mu\text{g/L}$

and the mean concentration in 2006 was 308 $\mu\text{g/L}$.

Data in Figure 3 are expressed on a log scale, and all statistical calculations in the Second Phase were done on the natural logarithms of the concentrations. As a result, the means in Figure 3 are geometric means instead of the more familiar arithmetic means. To calculate an arithmetic mean, one adds the values and divides the sum by the number of values.

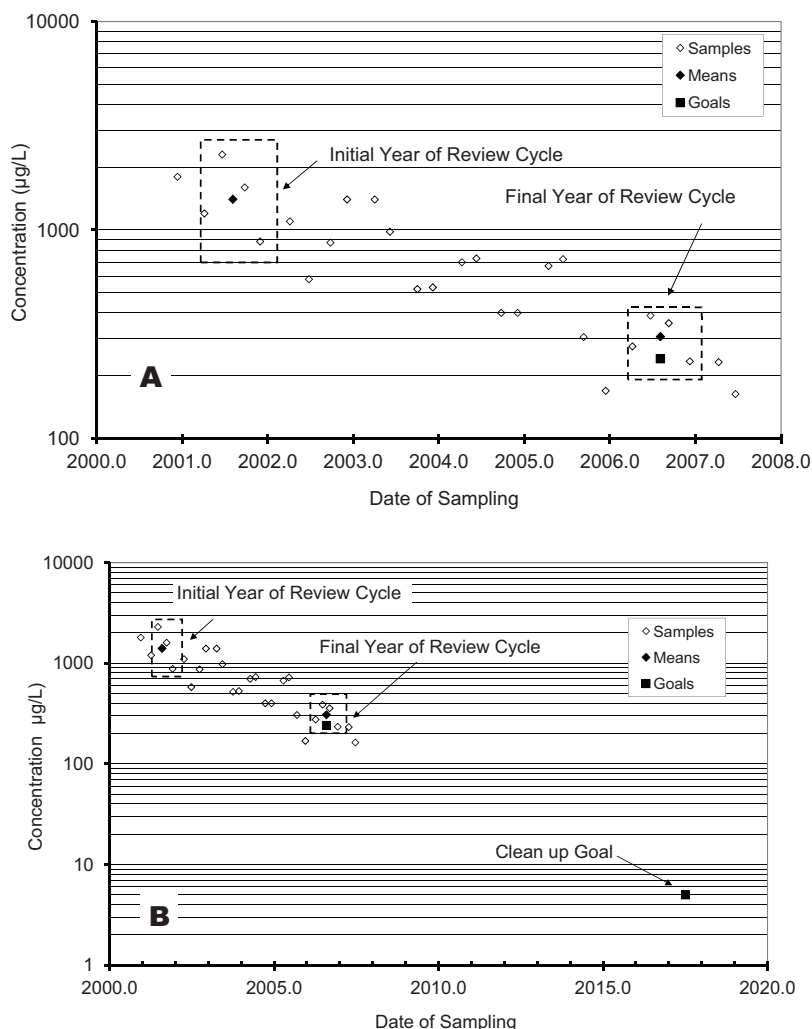


Figure 3. Comparison of the means for the first year of the review cycle (2001) and the final year of a five year review cycle (2006) to the clean up goals.

An arithmetic mean was calculated for the natural logarithms of the concentrations, and then the anti-natural logarithm was taken to express the geometric mean as a concentration. This is equivalent to multiplying the original values and taking the n^{th} root of the product, where n is the number of values. Note that for a set of positive values, the geometric mean is less than the arithmetic mean. The larger numbers have less influence on a geometric mean.

3.1. *Testing whether the reduction in concentration is statistically significant*

First, the data were evaluated to determine whether there was a statistically significant decline in the average concentration of contaminants over the time interval in the review cycle. Over the time period in the review cycle, was there any significant attenuation? The comparison is made using the t statistic for the difference of means. The theoretical background on the t statistic and details on using a spreadsheet to make the calculations are presented in Appendix B.

The t statistic allows the user to determine whether means are different from each other if the user is willing to accept some previously selected probability that the results of the statistical test may be in error. If the statistical test indicates that attenuation is significant at the selected probability of error, then it is appropriate to go forward with the remainder of the evaluation. If the test fails to indicate that attenuation is significant over the time interval in the review cycle, then it is not appropriate to interpret any apparent decline in concentrations over the time interval in the review cycle. Attenuation may have happened, but the attenuation is not documented by the monitoring data.

Imagine that the acceptable probability of error for the attenuation of concentrations of TCE as presented in Figure 3 was 10%. Appendix B presents detailed calculations that determine that the extent of attenuation between the mean of samples taken in 2001 and the mean

of samples taken in 2006 is in fact statistically significant with a probably of error of 10%, and it is appropriate to go to the next step in the evaluation.

3.2. *Testing whether the reduction in concentration is adequate to meet goals*

The next step is to determine whether the observed mean of the concentration data in the final year of the review cycle is consistent with meeting the clean up goal in the predetermined time interval. To do this, it is necessary to calculate a goal in the final year of the review cycle that is consistent with the predetermined ultimate long term goal for clean up. Because natural attenuation can be expected to follow a first order rate law, we will assume the same extent in attenuation (C_i/C_o) in each review cycle. Equation 2 can be rearranged to produce Equation 8.

$$\ln(C_i/C_o) = kt \quad 8$$

When the review cycles are the same length, the rate of attenuation ($\mu\text{g/L per year}$) decreases as the concentrations go down, but the extent of removal (C_i/C_o) in each review cycle stays the same.

The consequence of this assumption is illustrated in Equation 9, where for purposes of illustration there are six review cycles between the first year of the first review cycle and the date by which the clean up goals are to be attained. The product of the attenuation in each cycle will be the overall attenuation, where C_o is the concentration at the start of the first review cycle, and C_1 , C_2 , C_3 , C_4 and C_5 are the concentrations at the end of the current, second, third, fourth and fifth review cycles respectively, and C_g is the ultimate long term goal at the end of the sixth review cycle.

$$\left(\frac{C_1}{C_o}\right)\left(\frac{C_2}{C_1}\right)\left(\frac{C_3}{C_2}\right) \quad 9$$

$$\left(\frac{C_4}{C_3}\right)\left(\frac{C_5}{C_4}\right)\left(\frac{C_g}{C_5}\right) = \left(\frac{C_g}{C_o}\right)$$

$$\text{If } \left(\frac{C_1}{C_o}\right) = \left(\frac{C_2}{C_1}\right) = \left(\frac{C_3}{C_2}\right) = \left(\frac{C_4}{C_3}\right) = \left(\frac{C_5}{C_4}\right) = \left(\frac{C_g}{C_5}\right),$$

$$\text{then } \left(\frac{C_i}{C_o}\right)^6 = \left(\frac{C_g}{C_o}\right).$$

In general, if there are n review cycles between the first year of the first cycle and the goal, then

$$\left(\frac{C_i}{C_o}\right)^n = \left(\frac{C_g}{C_o}\right) \text{ and } \left(\frac{C_i}{C_o}\right) = \left(\frac{C_g}{C_o}\right)^{1/n}.$$

This final term can be rearranged to produce Equation 10, which calculates the interim concentration goal C_i from the initial concentration C_o , the final goal C_g , and the number of review cycles n .

$$C_i = C_o \left(\frac{C_g}{C_o}\right)^{1/n} \quad 10$$

For purposes of illustration, we will assume that the monitoring wells must reach the MCL for TCE by 2017, and that there will be five years in each review cycle. The initial year in the first review cycle started in 2001: the cleanup goal must be attained within 16 years of the initial year of the review cycle. There is time for $16/5 = 3.2$ review cycles (n) including the current cycle.

The concentration of TCE in the first sample taken in 2001 was 1200 $\mu\text{g/L}$. For purposes of illustration, we will assume that the clean up goal is the MCL for TCE, which is 5 $\mu\text{g/L}$. In the example, $1/n = 1/3.2$. Using these assumptions, Equation 10 becomes:

$$C_i = 1200 \left(\frac{5}{1200}\right)^{1/3.2} = 216 \mu\text{g/L}.$$

Panel A of Figure 4 compares the samples in the initial year of the first 5-year review cycle to the interim goals set using Equation 10 for the final year of the review cycle and to the final clean up goal. Panel B of Figure 4 compares the interim goals to the actual samples taken in 2006. The concentration of TCE in

the samples was higher than the concentration in the interim goals, raising the possibility that natural attenuation might not be adequate to attain the ultimate clean up goals. If the concentration of TCE were lower than interim goals, this fact would indicate that natural attenuation over the review cycle was adequate to meet the goals and no further statistical evaluation of the data would be necessary.

3.3. Estimating the probability that the reduction in concentration is not adequate to meet goals

The comparison in Panel B of Figure 4 does not take into account the natural variation in the data. If MNA were exactly adequate to meet the goal, but subject to random variation from one review cycle to the next, approximately one-half the time the samples would be greater than the concentration required to meet the goal, and one-half the time they would be less. In Panel B of Figure 4, the interim goal for each sample taken in 2006 was lower than the actual concentration, and the mean of the interim goals was lower than the mean of the actual samples, but the concentration in some of the samples taken in 2006 is less than the interim goals for other samples. It is not obvious from direct inspection whether the mean of the samples is so much greater than the mean of the goals as to be statistically significant. How much concern should we have that the concentrations in 2006 were not adequate to meet the clean up goal within the specified timeframe?

3.3.1. Establishing a decision criterion

The following decision criterion will be used to determine if attenuation is not adequate to meet the long term goal. **If the mean of the interim goals in the final year of this review cycle is less than the mean of the samples in the final year of this review cycle at some predetermined level of confidence, then attenuation in this review cycle is not adequate to attain the goal.** The null hypothesis in the comparison is “The mean of the interim goals is not less than the mean of the samples in the final year of the review cycle.” The interpretation

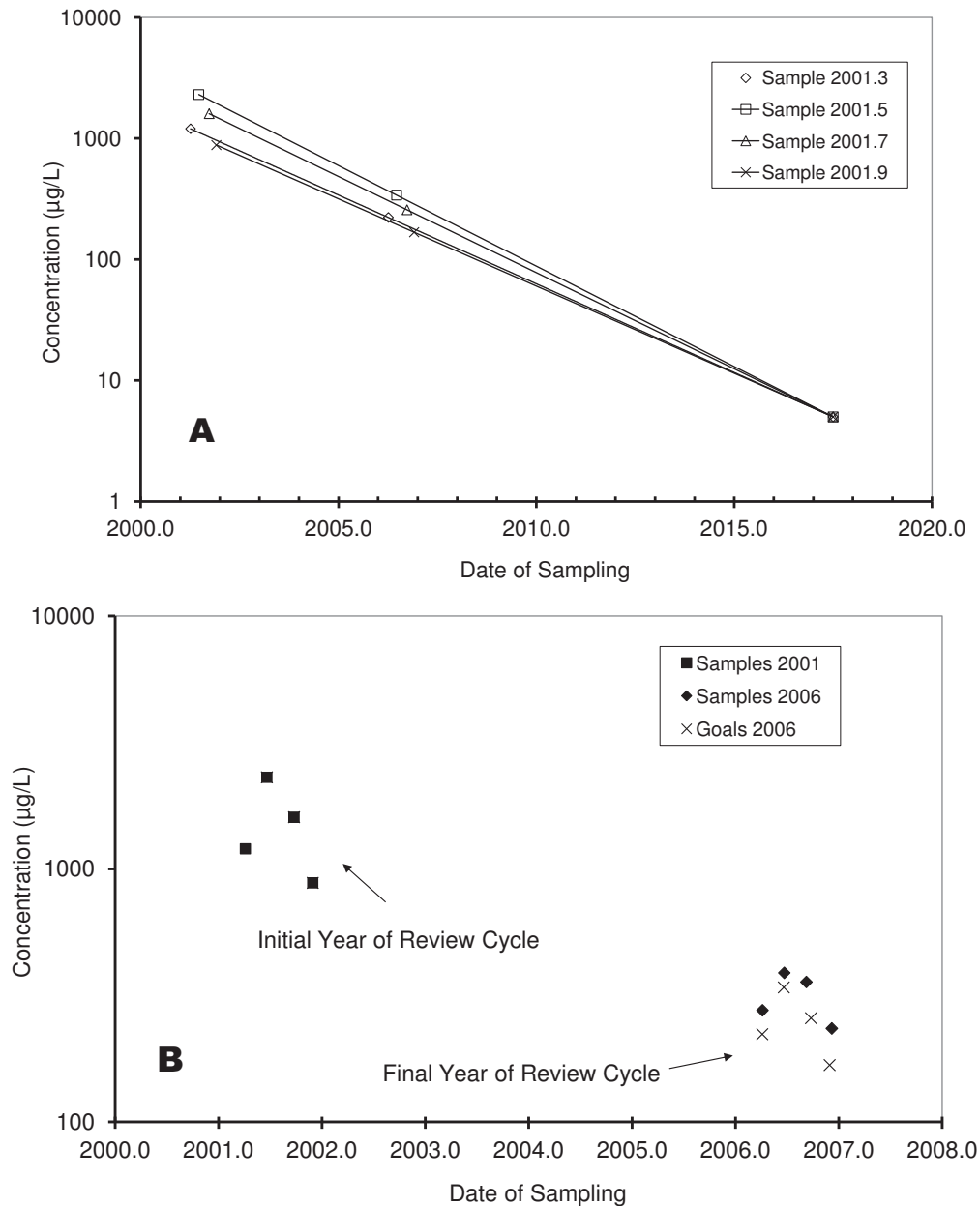


Figure 4. A comparison of the concentrations in samples in the initial and final year of the review cycle to the interim goals in the final year of the cycle (2006) to be adequate for the long term clean up goal (5 µg/L in 2017).

of the null hypothesis is “Attenuation in this review cycle is adequate to meet the goals.” If the null hypothesis is rejected, then the interpretation is “Attenuation in the review cycle is not adequate to meet the goals.”

3.3.2. Balancing the confidence in the test and the power of the test

Statisticians recognize two types of error in statistical comparisons. Type I error is a false positive. The statistical test rejects the null hypothesis when the null hypothesis is true. When a false positive result occurs, the statistical test indicated that the mean of the interim goals is less than the mean of the samples when

in fact the means are not different. According to the decision criterion, the statistical test would indicate that attenuation is not adequate to meet the goal, when in truth it is adequate. The probability of a Type I error is represented by α . The confidence in the test is described by $(1 - \alpha)$, usually expressed as a percentage.

Type II error is a false negative, in this case failing to reject the null hypothesis when in truth the null hypothesis is false. In this application the error would be to fail to indicate that natural attenuation was not adequate to meet the goal, when in truth it is not adequate. The probability of a Type II error is represented by β . The statistical power of the test to recognize when natural attenuation is not adequate is described by $(1 - \beta)$.

The null hypothesis (H_0) is- The rate of attenuation is adequate to attain the cleanup goal by the time specified.		
	H_0 is true	H_0 is false
Accept H_0	Correct Decision	Type II error (probability β)
Reject H_0	Type I error (probability α)	Correct Decision

For a given data set and statistical test, α and β stand in an inverse relationship to each other. As the value of α becomes lower (the test is more stringent, and confidence is higher), the value of β becomes higher and the test has less power to recognize when apparent attenuation is not adequate. The smaller the value of α , the less it is likely that the statistical analysis will warn that attenuation is not adequate to meet the long term goal.

Both types of error are important to an evaluation of natural attenuation; a value for α should not be selected based on some previous rule-of-thumb or default value. One who is trying to support the use of MNA might want to be able to recognize MNA when it is happening. This person would want to minimize Type I error, and would select small values of α and

corresponding large values of β . Alternatively, one who is concerned that attenuation is not adequate to meet the long term goal might want to be sure that the statistical test warns that attenuation is not adequate, when in fact, attenuation is not adequate. This person would want to minimize Type II error, and would select large values of α and corresponding small values of β .

One objective approach is to balance Type I and Type II error and to perform the test for values of α and β where $\alpha = \beta$. Appendix B describes in detail one process to determine these values. There are many site-specific issues and other factors that need to be considered while selecting the appropriate values for α and β . This document does not discuss the factors that might go into the decision. If site specific concerns make either Type I error or Type II more important, the ratios of α to β can be adjusted accordingly as described in Appendix B. With small data sets, balancing Type I and Type II error will result in relatively large values of both α and β .

3.3.3. Applying the decision criterion

The t -test for the difference of means will be used to determine if the mean of goals is less than the mean of the samples. Appendix B includes a short discussion of the assumptions behind the t -statistic, and provides instructions to set up a spreadsheet (EvalMNA) to make the calculations. As illustrated in Appendix B, EvalMNA first determines geometric means of the samples in the initial year (C_0) and final year (C_i) of the review cycle. In order to make the data more closely follow the underlying assumptions of the t -statistic, all the calculations are done using the natural logarithms of the concentrations and not the concentrations themselves. A mean calculated on the natural logarithms is equivalent to a geometric mean of the original data set, not an arithmetic mean of the original data set.

For the data presented in Panel B of Figure 3, these geometric means are 1404 $\mu\text{g/L}$ and

308 $\mu\text{g/L}$ respectively. EvalMNA then calculates the attenuation factor (C_i/C_o) that would be significant at various levels of the probability of error. At a value of (C_i/C_o) = 1.0, there is no attenuation. The attenuation of the means over the review cycle (C_i/C_o) = 0.219 was significantly different from one at a probability of error that was less than 0.0025.

Then EvalMNA compares the geometric mean of the interim goals to the geometric mean of the samples in the final year of the review cycle. The mean of the interim goals that would be necessary to be adequate for MNA was 241 $\mu\text{g/L}$. Panel B of Figure 3 compares the means of the samples to the mean of goals. For the data set in Panel B of Figure 3, when $\alpha = \beta$, then $\alpha = \beta = 0.26$ (See Section B.3 in Appendix B for details on how the values of α and β were determined for this data set). At this value of α , there is 26% chance of accepting that attenuation is not adequate to meet the goal, when in truth it is adequate. At this level of β , there is only a 74% chance of recognizing that attenuation is not adequate, when it is not adequate. Based on the decision criterion, the t -test for the difference of means indicated that attenuation of concentrations of TCE in the well during the review cycle was not adequate to meet the long term goal. See Section B.1.3. for details of the calculations.

Regression as an Alternative Phase Two Analysis

One potential criticism of the approach suggested for the second phase of analysis is that the comparison of the initial and final year of the review cycle does not take advantage of all the data that have been collected. The concern does merit consideration. Regression is designed to evaluate the entire range of the data, and is not specifically designed for the evaluation of the changes from the beginning to the end of a review cycle. However, a comparison of an interim goal to a statistical confidence belt on the regression line can be used to determine if the rate of attenuation is not adequate at some level of statistical confidence. If the regression line falls below the interim goal, the rate of attenuation is faster than the rate required to meet the goal, and there is no indication from the trend in concentrations that the progress of MNA is anything other than satisfactory. If the regression line falls above the interim goal, the rate of attenuation is too slow to attain the goal. However, this simple criterion ignores the uncertainty in the regression line. A better approach is to compare the interim goal to a statistical confidence belt on the regression line.

If regression is to be used to determine whether attenuation is adequate for the long term goal, the interim goal is compared to a line instead of a mean. The interim goal should represent the properties of the line, instead of the mean condition in the first year of the review cycle. The average properties of all the data in the regression are described by the midpoint of the regression line. The interim goal should be calculated by comparing the midpoint of the regression line to the final clean up goal, using Equation 10. If \bar{X} is the mean of the sampling dates in the regression, $X_{interim\ goal}$ is the last date in the regression, and $X_{final\ goal}$ is the date by which the concentration based goal is to be

attained, the value of n in Equation 10 would be determined following Equation 11:

$$n = (X_{interim\ goal} - \bar{X}) / (X_{final\ goal} - \bar{X}) \quad 11$$

The value of C_o in Equation 10 is \bar{Y} the mean of the values of Y in the regression, where Y is the natural logarithm of the concentrations of contaminants.

There may be negative consequences to a MNA remedy that is working too slowly, but no negative consequences if the remedy is working faster than necessary to attain the goals. For this reason, the statistical confidence belts should be one-tailed estimates on the regression line; all the probability for error in the regression should be assigned to the lower belts. These belts describe the rates of attenuation that are faster than the regression line. If the confidence belt is higher than the value of the interim goal, then at that level of confidence the rate of natural attenuation over the time period represented in the regression is too slow to attain the clean up goal by the time specified.

4.1. Establishing a decision criterion for regression

The decision criterion would be as follows. **If the interim goal lies below the confidence belt, then natural attenuation over the time interval included in the regression is not adequate to meet the goal.** Figure 5 compares the final clean up goal and the interim goal for the example data set to a linear regression on data collected in the years 2001 through 2006. The solid line in Figure 5 is the best estimate of the “true” line through the data. The regression line is slightly above the interim goal for the end of the monitoring period. In Figure 5, it is difficult to see the relationship between the regression line and the interim goal. Panel A of Figure 6 plots the same data on a time scale

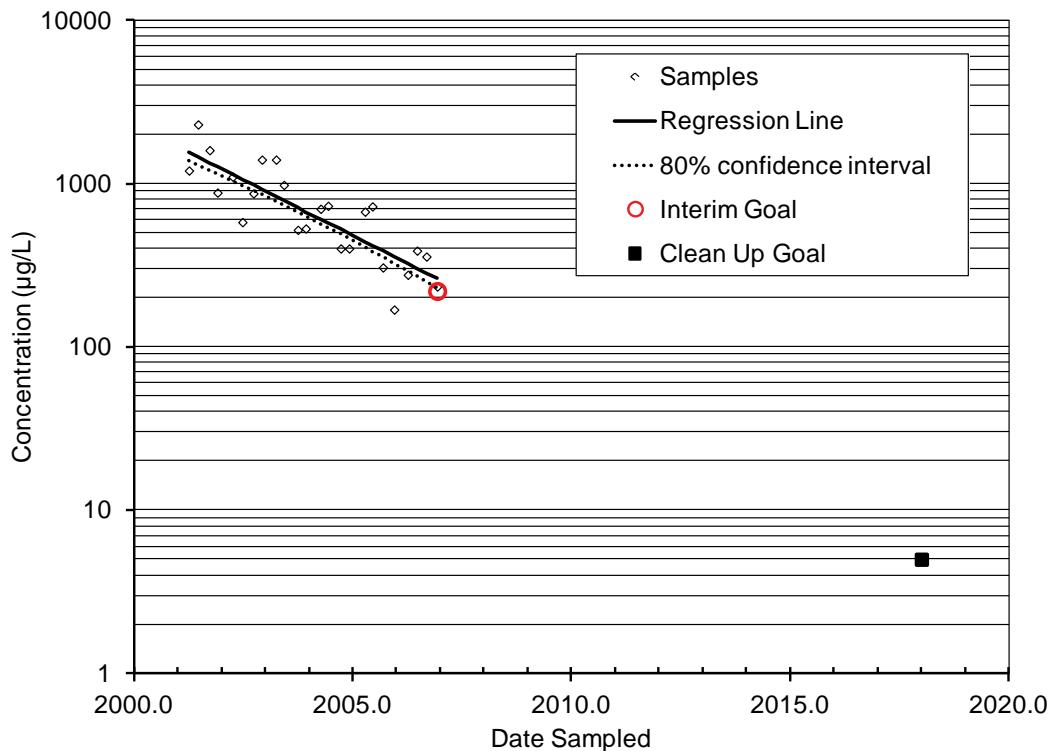


Figure 5. A comparison of the regression line through the natural logarithm of the concentrations of TCE, and the 80% confidence interval on the line to interim goal at the end of the review cycle (2007) and the long term clean up goal (5 µg/L in 2017).

of ten years. Note that the interim goal falls between the 80% and 95% confidence belts. At 80% confidence we can say that attenuation is not adequate to meet the long term goal, but we cannot say the same thing at 95% confidence.

If the value of α for the one-tailed confidence interval on the regression line is adjusted until the confidence belt runs through the interim goal, that value of α is 0.127 corresponding to a probability of error of 87.3%. In comparison, the t -test for difference of means determined that the confidence level that could support a determination that natural attenuation was not adequate fell somewhere between 85% and 90% (In Appendix B, compare cell X15 to cells AA16 and AA17 in Figure B6 and cells AB16 and AB17 in Figure B6.). If values of α are adjusted in EvalMNA to determine the value where the prediction changes from *No evidence that attenuation is not adequate* to *Attenuation is not adequate to attain goal*, that value of

α is near 0.125 corresponding to a probability of error of 87.5%. The results of the two approaches were similar for the example data set.

Figures 5 and 6 were created in EXCEL. Equations 10 and 11 were used to calculate a value of the interim goal. The regression line and the confidence belts on the regression line were calculated following the standard formulas for linear regression. The EXCEL file used to create Figure 5 and 6 can be used as a template to calculate an interim goal and compare the interim goal to confidence belts from another monitoring record. To use the file as a template, download the file from <http://www.epa.gov/nrmrl/gwerd/csmos/models/RegressionMNA.xls>. Open the tab **Data and Calculations** and follow the instructions to copy the data from a new monitoring record over the data in the template.

4.2. Regression with only one sample each year

McHugh et al. (2011) evaluated three large monitoring records and found that there was a characteristic time-independent variability associated with subsequent samples from the same well. For most wells the long-term trend in attenuation could not be distinguished from the

time-independent variability unless the samples were separated in time by 320 to 400 days. Although this relationship was first documented in a formal manner in 2011, many site managers and regulatory staff have had an intuitive understanding of this interaction for years. As a result, many data sets only have one sample

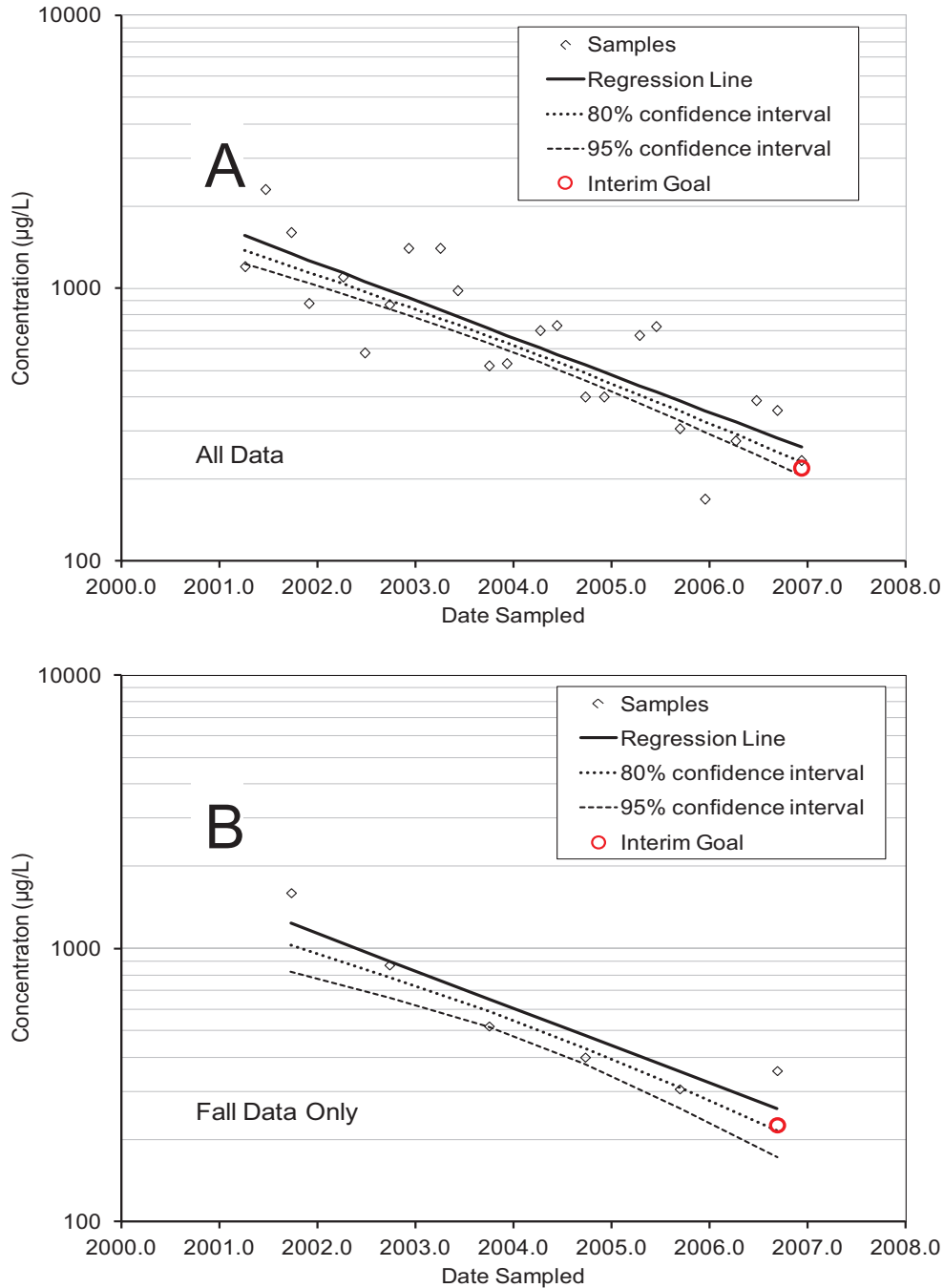


Figure 6. Effect of number of samples on the confidence belts on regression lines fit through the data.

in each year and the alternative approach using regression is the only approach available.

This situation is illustrated in Panel B of Figure 6. The data set in Panel A was collapsed by selecting only the fall sample for each year of the review cycle. Because there were fewer data in the regression, the confidence belts in Panel B are wider than in Panel A. The interim goal is above the 80% confidence belt in the regression in Panel B. Contrary to the case for the regression with twenty-four samples, the regression with six samples failed to indicate that attenuation was not adequate for the goal at 80% confidence. This illustrates the potential problem of performing statistical data analysis on smaller data sets. Smaller data sets have less capacity to recognize when attenuation is not adequate to attain the goal.

Putting the Statistical Analysis into a Geohydrological Framework

Any statistical analysis can be no better than the monitoring data. Statistics cannot substitute for an inadequate monitoring record. The statistical approach described here can only serve its purpose when it is applied to a system of monitoring wells that is adequate to describe the plume of contamination. Consult *Performance Monitoring of MNA Remedies* (Pope et al., 2004) for U.S. EPA recommendations concerning the design of a network of monitoring wells, and a process for developing site-specific monitoring objectives and performance criteria for MNA. Statistics can only provide a reliable interpretation when they are applied to data from wells that are representative of the groundwater conditions and adequately characterize the plume vertically and horizontally. Pope et al. (2004) put emphasis on the need for a monitoring network that can track the plume of contamination in three dimensions.

The example used to demonstrate the methodologies outlined in this paper focuses on a single well, but the clean-up goals must be met throughout the plume. The methodologies described in this approach should be applied to the entire monitoring network and not at only a few selected wells. Similarly, degradation rates may vary in different portions of the plume and over time, so the monitoring well providing the least favorable evaluation of MNA progress may not be the same well from one evaluation cycle to the next. Further, those applying the methods described in this paper should recognize that large scale environmental conditions may affect the results of any MNA evaluation. For example Section 3 uses data from the first and last year of a CERCLA 5-Year Review cycle to evaluate the progress of MNA and predict whether cleanup goals will be attained on time. This approach might overestimate

or underestimate the effectiveness of MNA if the site is impacted by either severe drought or flooding during the first or final year of a 5-Year Review cycle. This concern diminishes as the number of observations (samples) and the number of review cycles becomes larger over time, but the impact of factors external from the site should not be ignored especially when the database is small. Statistical methods can not compensate for bad data, inadequate data, an inaccurate concept model, or insufficient experience in the interpretation of groundwater flow systems and contaminant migration.

5.1. Monitoring to document trends in attenuation in three dimensional space

As described in Section 2.5 of *Performance Monitoring of MNA Remedies* (Pope et al., 2004), the direction of ground water flow can change from season to season and from year to year. A subtle shift in the direction of ground water flow can move a plume away from a monitoring well, giving the false impression that the concentrations of contaminants are attenuating. Pope et al. (2004) offer suggestions for dealing with these changes. If monitoring wells are arranged into transects that are aligned perpendicular to ground water flow, it is often possible to recognize when a plume has shifted, as opposed to being attenuated. As the plume moves away from one well it will move toward an adjacent well in the transect. Over time, plumes can shift up and down in aquifers as well as side to side. To deal with this possibility, well clusters can be installed at different elevations in the monitoring well transect.

At the site illustrated in Figure 1, each monitoring location is a cluster of wells which sample water from different portions of the aquifer.

Figure 1 presents the data from the most contaminated well in the cluster at each monitoring location. Table 1 presents data on the vertical distribution of contamination in the monitoring well transect. At the most contaminated location, there was significant variation in concentrations with depth; the highest concentrations of contaminants were at the shallowest depth interval.

The monitoring wells in the transect capture flow lines of contaminated water in the plume as well as clean or much cleaner ground water that surrounds the plume in three-dimensional space. To the extent that the monitoring wells provide a comprehensive sample of flow in the plume, the trend analysis of concentrations in the monitoring wells will provide a comprehensive description of attenuation of contaminants. Because concentration trends in monitoring wells may vary independently of initial contaminant concentrations, it is important to evaluate concentration trends across the entire contaminant plume.

5.2. Evaluation of whether attenuation is adequate at a site

A remedy is put in place to protect a valued resource from contamination or to restore a valued resource, an aquifer, to beneficial use. If necessary, natural attenuation can be replaced with some form of active remediation that is designed to treat the entire plume, or active remediation can be applied to those portions of the contaminant plume that are not performing as required in the ROD. *Performance Monitoring of MNA Remedies* (Pope et al., 2004) identifies several decisions that can be made as part of a performance review: continue monitoring without change, modify the monitoring program, modify institutional controls, implement a contingency or alternative remedy, and terminate performance monitoring. Pope et al. (2004) identify six criteria that may trigger a decision. Three of the criteria address trends in concentrations. These criteria are:

- 1) contaminant concentrations in soil or ground water at specified locations

exhibit an increasing trend not originally predicted during remedy selection,

- 2) near-source wells exhibit large concentration increases indicative of a new or renewed release, and
- 3) contaminant concentrations are not decreasing at a sufficiently rapid rate to meet the remedial objectives. The approach presented in this Report is designed to evaluate the third criterion.

As discussed in Section 2.6.1.2 of Pope et al. (2004), most reviews of the progress on natural attenuation attempt a comprehensive assessment of the performance of the plume as a whole. The assessment may use Thiessen polygons to associate concentrations in individual wells with a volume of water, multiply concentration by volume to estimate the total mass of contaminant in the polygon sampled by the well, and then add the masses of all the polygons to estimate the overall mass of contaminant in the plume at any given time (Dupont et al., 1998a, 1998b). The overall rate of natural attenuation is calculated by comparing the rate of reduction of total mass over time. Other assessments simply tally the number of wells where concentrations are declining, stable, or increasing.

Under most regulatory programs, the site will not be considered clean until the entire site is clean. Consequently, MNA can be demonstrated to be an acceptable remedy when the methods described in this paper show that all of the monitoring wells can be expected to achieve the site-specific clean up goals for each contaminant within an acceptable timeframe.

5.2.1. Evaluation of the site as a whole

Any large data set will likely have some problematic wells that:

- 1) are not meeting expectations, or
- 2) show an increase in concentrations over the review cycle, or
- 3) have so much variation in concentration over time that it is impossible to discern a trend.

A comparison of any problematic well to the most contaminated wells at the end of the review cycle will put the problematic well in context. **If the final concentration in a problematic well at the end of the review cycle is low with respect to the most contaminated wells at the site, then that particular problematic well, alone, does not put the site at risk for not attaining the final goal, at least not in the review cycle under consideration.** If the problematic well is one of the most contaminated wells at the end of the review cycle, then the problematic well indicates a risk that the remedy for the site will not attain the final goal. A decision is required as to whether an additional monitoring cycle is appropriate or whether a more active remedial measure should be implemented in all or a portion of the plume.

The second phase of analysis facilitates these comparisons. The interim goal that is set for the well that has the highest initial concentration of a particular contaminant can be considered the site-wide interim goal for that contaminant for the review cycle. If no well has a concentration of that particular contaminant at the end of the review cycle that is above the site-wide interim goal, then attenuation across the entire site can be considered adequate to attain the final clean up goal. If a different well exceeds the site-wide interim goal at the end of the review cycle, a re-evaluation of MNA as a potential remedy should be conducted. The Conceptual Site Model (CSM) should be reevaluated in light of the new information regarding the distribution of contamination. The possibility of additional source areas or residual source material should be evaluated. The possibility that groundwater flow directions have changed, due to pumping or simply a re-interpretation of the water level data, should be evaluated. If the exceedance of the interim goal is small, if revised estimates of the final clean-up time do not conflict with the conditions in the ROD or with other expectation for use of the resource, and if the explanations (CSM, sources, flow direction changes due to pumping) are reason-

able and manageable, then monitoring for an additional cycle may be appropriate.

If the contaminant concentration in the well newly observed to exceed the interim goal would jeopardize the cleanup time expectations of the ROD or the reuse expectations for the site, and if additional source areas or residual source material are not found and removed, then additional monitoring for another cycle may not be justified and more active remediation should be implemented as described in Section 5.2. This situation is described in the next few paragraphs using data from the example site. A method for quantifying the uncertainty related to the decision to continue MNA for another review cycle is described in Section 5.2.2.

Table 2 illustrates these well-to-well comparisons for concentrations of TCE, *cis*-DCE, and vinyl chloride in the three cluster wells in location MW 3 in Figure 1. For each compound in each monitoring well, the EvalMNA spreadsheet was used to estimate the geometric means of concentrations at the beginning and end of the review cycle, as well as the interim goal. The mean concentration in the final year was compared to the interim goal to determine if attenuation was adequate to meet the long term goal. If the calculations in EvalMNA indicated that attenuation was not adequate, then Table 2 presents the probability that the calculations were in error, and attenuation truly was adequate. If the probability of error is small, that is a strong indication that attenuation was not adequate.

The highest initial concentration of TCE was found in well MW-3C, and the interim goal for MW-3C (582 µg/L) can be considered the interim goal for TCE for the entire site. Fortunately, attenuation of TCE in MW-3C was adequate to attain the long term goal. The final mean for TCE in MW-3C was 259 µg/L. In contrast, attenuation of TCE in well MW-3B was not adequate to meet the long term goal; and there is only a 15% chance that attenuation would truly reach the goal. However, the final

Table 2. Second phase of analysis. Evaluation of whether the reductions in concentrations of TCE, *cis*-DCE and Vinyl Chloride in monitoring over the review cycle will not meet the goals by the specified date.

Well	Compound	Mean Conc. in 2001 C_o	Mean Conc. in 2006 C_i	Interim Goal for 2006 C_g	Is attenuation adequate to meet long term goals?*	α^{**}
		$\mu\text{g/L}$	$\mu\text{g/L}$	$\mu\text{g/L}$		
MW-3C	TCE	5055	259	582	Yes	
MW-3B	TCE	1404	308	241	No	<0.15
MW-3A	TCE	Below MCL				
MW-3C	<i>cis</i> -DCE	1648	35	614	Yes	
MW-3B	<i>cis</i> -DCE	2319	781	777	No	>0.4
MW-3A	<i>cis</i> -DCE	1282	493	517	Yes	
MW-3C	Vinyl Chloride	384	10	74	Yes	
MW-3B	Vinyl Chloride	1652	559	203	No	<0.0025
MW-3A	Vinyl Chloride	1096	604	153	No	<0.0025

* If C_i (the geometric mean for 2006) < C_g (the interim goal for 2006), then attenuation over the review cycle is adequate to attain the clean up by the date specified in the ROD.

**The value of α is the probability that the statistical test in EvalMNA will indicate that attenuation is not adequate when in truth it is adequate.

concentration of TCE in MW-3B (308 $\mu\text{g/L}$) was less than the interim goal for the entire site (582 $\mu\text{g/L}$). Despite the fact that attenuation of TCE in well ME-3B was not adequate to attain the long term goal, attenuation across the entire site was adequate in this review cycle.

Well MW-3B had the highest concentrations of TCE at the site in 2006, and the interim goal for MW-3B becomes the interim goal for the entire site in the next review cycle. If attenuation of TCE in MW-3B continues along the same course, then attenuation of TCE will be revealed to be inadequate in the next review cycle. In this situation, there may be justification to continue monitoring for another review cycle to see if the problematic behavior was a statistical aberration, or whether it is a consistent behavior for that well. As described

below, the results from comparisons for other contaminants may impact the decision whether a decision about this problem well should be postponed for another review cycle.

Monitoring well MW-3B had the highest concentrations of *cis*-DCE at both the beginning and end of the review cycle. A comparison of the mean concentration at the end of the review cycle to the interim goal indicates that attenuation was not adequate, but the mean concentration almost attained the interim goal, and the probability is high that the calculations are in error. There would be some justification to wait another review cycle to gather more compelling evidence that attenuation of *cis*-DCE in well MW-3B was not adequate before a decision was made to alter the ROD.

Along with Monitoring well MW-3A, Monitoring well MW-3B also had the highest concentrations of vinyl chloride at the beginning and end of the review cycle, and the second phase of analysis indicates that attenuation is not adequate to meet the long term goal. In contrast to the situation with *cis*-DCE, the probability of error is very low. There is little chance that the attenuation of concentrations of vinyl chloride over the current review cycle is adequate to meet the goals, and the behavior of vinyl chloride deserves attention. The behavior of vinyl chloride will be discussed further in the next section.

5.2.2. Evaluation of individual wells

If the analysis in the second phase indicates that attenuation over the review cycle is not adequate to attain the clean up goal, the next step is to evaluate the potential consequences. The simplest and most straightforward way to make that evaluation is to compare the date

when concentrations might achieve the clean up goal to the date specified in the ROD. Again, the process will be illustrated with the example data set. Equation 4 was used to estimate the date to attain the MCL for each contaminant in each well.

Table 3 compares the maximum concentrations and rate constants for attenuation to the date when the goal might be attained. All three contaminants might fail to meet the goal by 2017. Projected concentrations of TCE miss the goal by less than two years while projected concentrations of *cis*-DCE miss the goal by a little over six years and projected concentrations of vinyl chloride miss the goal by a little over thirty one years. Concentrations of TCE are projected to be below the MCL by the end of the final review cycle, while concentrations of *cis*-DCE and vinyl chloride above the MCL are expected to persist into subsequent review cycles.

Table 3. Using the first phase of analysis to project when concentrations of TCE, *cis*-DCE and Vinyl Chloride in monitoring wells will attain the clean up goal.

Well	Compound	Maximum Conc.	Rate Constant for Attenuation	Rate Constant at 90% Confidence	Date to Attain MCL
		µg/L	per year	per year	years
Target Date to Attain the Clean up Goal					2017
End of Fourth Review Cycle					2021
MW-3C	TCE	8600	0.588	>0.523	2013.5
MW-3B	TCE	2300	0.326	>0.281	2018.9
MW-3C	<i>cis</i> -DCE	3300	0.563	>0.380	2006.2
MW-3B	<i>cis</i> -DCE	3200	0.153	>0.106	2023.3
MW-3A	<i>cis</i> -DCE	1900	0.154	>0.105	2018.6
MW-3C	Vinyl Chloride	1200	0.715	>0.601	2009.3
MW-3B	Vinyl Chloride	2200	0.206	>0.152	2032.7
MW-3A	Vinyl Chloride	1500	0.130	>0.0832	2048.5

5.2.3. Uncertainty in estimates of time to attain cleanup goal.

When Equation 4 is used to estimate the date to attain the clean up goal, there is uncertainty in the estimate, and the uncertainty should be considered in the evaluation. Figure 7 compares the trend in concentrations of vinyl chloride in monitoring well MW-3A to the MCL for vinyl chloride. In Figure 7, the left hand Y-axis scales the data as the natural logarithm of vinyl chloride concentration, and the Y-axis on the right hand side scales the actual concentrations of vinyl chloride. A regression was used to predict the date when the concentration should equal the MCL. The uncertainty is evaluated by comparing the date when the confidence belt on the regression line reached the MCL to the date when the regression line reached the MCL.

The MCL for vinyl chloride is 2 µg/L, and the natural logarithm of 2 is 0.69. The regression line crosses the MCL at the projected date of 2048. The confidence belts were calculated for $\alpha=0.40$. There is a 20% chance that the true line runs above the upper confidence belt and

a 20% chance that the true line runs below the lower confidence belt. The lower confidence belt crosses the MCL in 2040. There is only a 20% chance that concentrations will attain the MCL before 2040.

Similarly, the upper confidence belt in Figure 7 crosses the MCL in 2063. There is a 20% chance of attaining the goal sometime after 2063, which means there is an 80% chance of attaining the goal before 2063. The value of α can be adjusted based on site specific concerns to identify an appropriate target date to attain the cleanup goal.

5.2.4. Interpretation of projections

The interpretation of contaminant concentration trends in an evaluation of the progress of natural attenuation depends on two separate evaluations. The first is a simple comparison of the concentrations at the end of the review cycle (such as a 5-Year Review) to the interim concentration goal to determine whether attenuation during the review cycle is adequate to attain the final clean up goal by the date

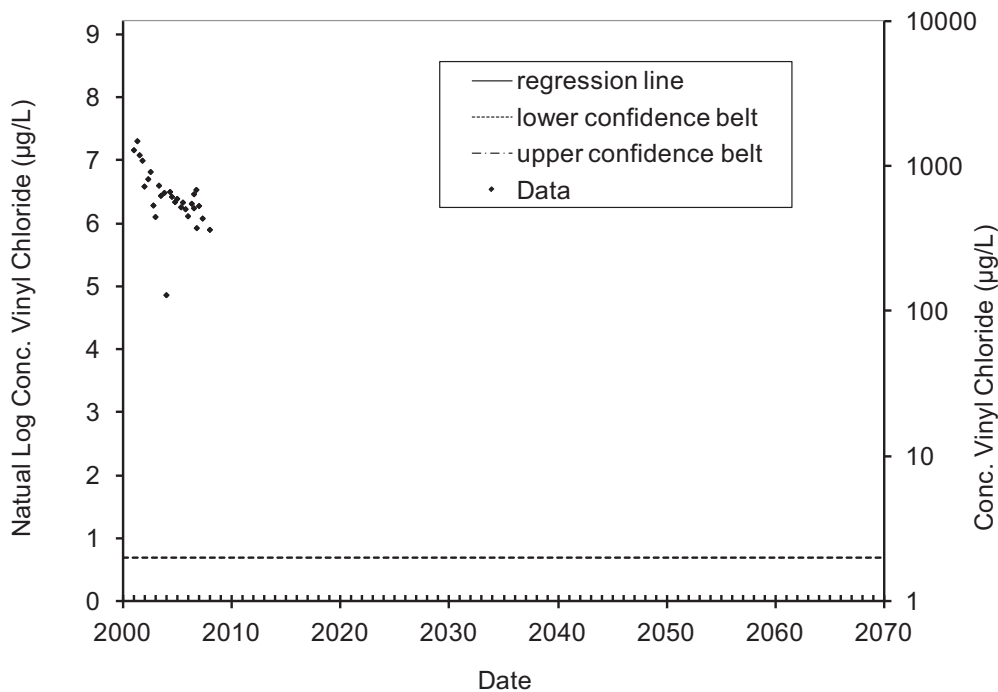


Figure 7. Projected 60% confidence belts on the regression of the natural logarithm of concentration of vinyl chloride in MW-3A on date.

specified in the ROD. If the concentrations can be expected to meet the clean up goal on or before the specified date, there is no need for further action until the next review.

If the comparison shows that attenuation cannot be expected to meet the clean up goal, the next step is to determine the statistical validity of that result. To make this determination, it is necessary to select an acceptable probability of error α to use with the statistical test (the confidence level of the test will be equal to $(1-\alpha)$). There are two approaches for obtaining α . In one approach, the value of α is selected (or negotiated) based on a willingness to modify the ROD and implement an active remedy. The other approach recognizes that α (the probability of Type I error) and β (the probability of Type II error) are inversely linked. [See Section 3.3.2 for a discussion of Type I and Type II error.] In the second approach, the ratio of β to α is selected. This may also involve negotiations between the responsible parties and regulatory agencies. Appendix B describes one process to determine these values.

To complete the evaluation, the selected level of α is used in the statistical tests built into EvalMNA. There can be two results: either the monitoring data indicate that attenuation is not adequate to meet the goal, or the data cannot provide evidence that attenuation is not adequate at the selected level of confidence (the statistical null hypothesis).

If the evaluation of the data indicates that attenuation is not adequate to meet the goal at the selected level of statistical confidence, this provides documentation that there is a problem with monitored natural attenuation as a remedy, and some sort of active corrective action may be indicated. If the data does not provide evidence that attenuation is not adequate, there is no basis to require additional active correction action at this time. When the data does not provide evidence that attenuation is not adequate, one cannot necessarily presume that adequate attenuation is in fact occurring. There is little choice except to continue to collect and

evaluate monitoring data. Even though natural attenuation appears to be adequate to achieve the remedial goals, the site is not yet clean and subsequent data may demonstrate inadequate attenuation.

Once confidence in the contaminant concentration trends and rates of attenuation are established, collecting data over a longer time interval may be appropriate and sufficient to demonstrate achievement of the objectives of the following review cycle(s). If the true rate of attenuation is slower than the rate needed to meet the clean up goal, the discrepancy between the samples and the interim goals will grow larger over time and be easier to discern. When the data are highly variable and there are too few samples, the statistical power of the test described in this paper will be low; so increasing the interval between samples (decreasing sample frequency) may not be appropriate. The monitoring strategy should be evaluated. If an acceptable probability of error (α) for use with the statistical tests cannot be achieved with the current sample frequency, perhaps samples for performance monitoring should be collected more frequently than was originally anticipated.

5.2.5. Transformation products from natural attenuation may be a special case

The concentration of vinyl chloride is controlled by the rate of production of vinyl chloride from *cis*-DCE as well as the rate of degradation of vinyl chloride to ethylene. The rate of attenuation is a composite rate which includes the rate of production of the degradation product from the parent compound. In the example used in this document the actual rate of removal is higher than the compound specific rate of attenuation because the input of vinyl chloride from degradation of *cis*-DCE offsets some of the degradation of vinyl chloride to ethylene, resulting in an apparent degradation rate which is less than would be attained if there was no *cis*-DCE. As the concentrations of TCE and *cis*-DCE decline over time, the rate of attenuation of *cis*-DCE and vinyl chloride can be expected to increase as the mass of parent

compound is exhausted. But with the data available in early review cycles, a projection of the regression line may tend to overestimate the time required for transformation products such as vinyl chloride to attain the clean up goal. As a consequence, concentrations of *cis*-DCE and vinyl chloride may attain their MCL at some time before the date projected by Equation 4.

This pattern of a transient accumulation of the degradation product depends on an active and ongoing mechanism for further degradation of the product. If there is no mechanism for further degradation of the product, the product (such as vinyl chloride from *cis*-DCE) will simply accumulate. To understand if this pattern of transient accumulation applies to a particular site, it is necessary to know if the degradation product is being degraded at the current time.

The most direct line-of-evidence for degradation of a compound is the production of the degradation product such as vinyl chloride from *cis*-DCE and ethylene (or ethane) from vinyl chloride. There are other lines of evidence that can substantiate degradation of organic compounds. A variety of molecular biological tools are available to document the density of microorganisms in ground water that can degrade contaminants (Azadpour-Keeley et al., 2009). These tools are particularly useful to understand the natural attenuation of PCE and TCE to ethylene or ethane. The only organisms that are known to degrade *cis*-DCE to vinyl chloride or to degrade vinyl chloride to ethylene belong to the *Dehalococcoides* group of bacteria. This group of bacteria can easily be recognized and enumerated in ground water with commercially available molecular biological tools (Lu et al., 2006). Often, the ratio of stable isotopes of carbon, hydrogen, and chlorine will change in a predictable fashion as compounds are degraded in groundwater, and changes in the ratio of isotopes in the organic compound as water moves along a flowpath can be used to document degradation (Hunkeler et al. 2008).

There may be justification for making a distinction between a primary contaminant and its transformation products. In the example data, TCE was the material that was released, and there is a possibility that TCE remains in the aquifer as a non aqueous phase liquid (NAPL) which could act as a continuing source of TCE contamination in ground water. If attenuation of the primary contaminant is not adequate to meet the goal, some sort of active corrective action may be indicated.

Often, the only plausible source of the transformation product in ground water is degradation of the primary contaminant. If the attenuation of the transformation products is not adequate to meet the goal, there may be justification to continue monitoring to see if the rate of attenuation of the transformation products will increase as the concentration of the primary contaminant decreases.

5.2.6. Rates of attenuation reported in the literature

One approach to evaluate and validate rate constants that are extracted from data at a particular site is to compare those constants to constants extracted by others at other sites. Farhat et al. (2004) analyzed temporal records and extracted rate constants for petroleum hydrocarbons from 366 sites and for chlorinated solvents at 40 sites. Newell et al. (2006) reported temporal records for concentrations of PCE, TCE, DCE or TCA in monitoring wells: 52 records were available from 23 sites. Kampbell et al. (2000) reported temporal records for concentrations of benzene, toluene, ethylbenzene and total xylenes at five large fuel spills. Peargin (2001) reported temporal records for MTBE in 22 wells, benzene in 39 wells, and xylene in 34 wells that were screened across a smear zone at fuel spill sites. Wilson et al. (2005) reported rate constants from six MTBE spill sites. These rates are compiled in Table 4. In general, the rate constants for attenuation in the absence of active remediation were in the order of 0.1 to 1.0 per year. The rates extracted in this document from the example data for attenuation of

Table 4. Typical rates of attenuation of concentrations over time in monitoring wells.

Chemical	Median Rate	Maximum Rate	Number of Sites	Reference
	per year			
Benzene	0.22	0.41	5	Kampbell et al. (2000)
Benzene	0.12		39	Peargin (2001)
Benzene*	0.22	4.49	359	Farhat et al. (2004)
Toluene	0.31	0.44	5	Kampbell et al. (2000)
Toluene*	0.41	6.42	89	Farhat et al. (2004)
Ethylbenzene	0.087	0.41	5	Kampbell et al. (2000)
Ethylbenzene*	0.18	2.96	90	Farhat et al. (2004)
Xylenes	0.093	0.43	5	Kampbell et al. (2000)
Xylenes	0.07		34	Peargin (2001)
Xylenes	0.25	7.10	89	Farhat et al. (2004)
PCE	0.23	0.97	9	Newell et al. (2006)
PCE*	0.11	1.84	32	Farhat et al. (2004)
TCE	0.11	0.60	13	Newell et al. (2006)
TCE*	0.15	2.42	37	Farhat et al. (2004)
<i>cis</i> -DCE	0.16	0.17	2	Newell et al. (2006)
<i>cis</i> -DCE*	0.62	1.75	11	Farhat et al. (2004)
Vinyl Chloride*	0.07	1.44	18	Farhat et al. (2004)
1,1,1-TCA	0.34	0.61	6	Newell et al. (2006)
1,1,1-TCA*	0.15	1.50	23	Farhat et al. (2004)
1,2-DCA*	0.14	1.25	13	Farhat et al. (2004)
MTBE	0.04		22	Peargin (2001)
MTBE	0.23	0.70	6	Wilson et al. (2005)
MTBE*	0.08	6.9	78	Farhat et al. (2004)

* Rate constants for well at each site with the highest concentration.

TCE, *cis*-DCE and vinyl chloride ranged from 0.1 to 0.7 per year (Table 3). These rates are consistent with rates extracted by Newell et al. (2006) and Farhat et al. (2004) at other sites (Table 4).

The range of rate constants reported in Table 4 is narrow. The lower bound on the reported rate constants may be caused by an inability to resolve rate constants that are much below 0.1 per year with current approaches to monitoring. Of the 52 temporal records reported by Newell et al. (2006) for chlorinated solvents, attenuation in 28 of the records was not statistically significant at 95% confidence. As described by

Newell et al. (2002), the rate of natural attenuation in a monitoring well more closely tracks the rate of weathering of the source of contamination as opposed to the rate of natural degradation of the contaminant once it enters ground water. The upper bound on the rate constants may be limited by the rate of mass transfer of contaminant from the residual source of contamination to ground water or soil gas.

5.2.7. Dealing with data quality issues

When the progress of natural attenuation is substantial and concentrations approach the Maximum Contaminant Level (MCL) that is allowed in drinking water, the data set will

often include samples with concentrations that are below the method detection limit (MDL) or the reporting limit (RL). If the data are evaluated by regression analysis and the samples with concentrations below the MDL or RL occur after the mean date in the regression, include the sample and use the MDL or RL as value for the sample. This will respect the number of degrees of freedom in the data set, but will bias the regression line and confidence bands to project concentrations that are higher than the true concentration. This biased analysis is conservative in that it will indicate that there is less attenuation than the true extent of attenuation. If this biased analysis indicates that attenuation is still adequate to attain the goal, there is no concern with the biased analysis. If the biased analysis indicates that attenuation is not adequate to attain the goal, disregard the analysis and refer the statistical analysis to a professional statistician who will apply methods that are appropriate for data sets that contain censored data. If the samples with concentrations below the MDL or RL occur before the mean date in the regression, refer the statistical analysis to a professional statistician.

If the data are evaluated by comparing the mean of samples in the final year of the monitoring period to the mean of interim goals, and the samples with concentrations below the MDL or RL occur in the final year, conduct a biased comparison using the MDL or RL as the value for the sample. If this biased analysis indicates that attenuation is adequate to attain the goal, there is no concern with the biased analysis. If samples with concentrations below the MDL or RL occur in the initial year used to establish the interim goals, abandon the approach of comparing means in the final year of the monitoring period and analyze the data set using regression analysis.

If the value is flagged with an EPA Data Qualifier “J” indicating that it is an estimated value, use the value in the statistical analysis. If possible, obtain another water sample and

analyze it using a protocol that is less likely to produce a flag.

Suggestions and Recommendations

Follow the process outlined in *Performance Monitoring of MNA Remedies* (Pope et al., 2004) to develop site-specific monitoring objectives and performance criteria for MNA. Conduct the performance review with those objectives and criteria in mind.

Prior to a review of the performance of MNA, analyze the monitoring record to reveal:

- 1) the attenuation in concentration of each contaminant in each monitoring well over the review cycle (C_i/C_o); and
- 2) determine whether the attenuation is adequate to attain the long term goal by the date specified in the decision documents; or alternatively,
- 3) estimate the probability that attenuation is not adequate to attain the goal.

If the concentrations in a well are not expected to attain the clean up goal by the specified time, associate a level of statistical confidence and statistical power with that determination. When the confidence in a statistical test and the power of the test are high, the results of the test are more compelling, and provide more justification to initiate active (more aggressive) clean up actions rather than MNA.

Identify those wells that are more likely to prevent a site from attaining the goals for MNA. Then identify and assign priority to those areas of the aquifer that are the best candidates for focused active remediation. Re-evaluate progress toward clean-up and site closure at the end of the next review cycle to determine the impact of the focused active remediation efforts.

References

- Azadpour-Keeley, A., M.J. Barcelona, K. Duncan and J.M. Sufflita. 2009. The Use of Molecular and Genomic Techniques Applied to Microbial Diversity, Community Structure, and Activities at DNAPL and Metal Contaminated Sites. EPA/600/R-09/103. Available: <http://www.epa.gov/nrmrl/pubs/600r09103/600r09103.pdf>.
- Dupont, R.R., D.L. Sorensen, M. Kemblowski, M. Bertleson, D. McGinnis, I. Kamil and Y. Ma. 1998a. Monitoring and Assessment of In-Situ Biocontainment of Petroleum Contaminated Ground-Water Plumes. EPA/600/R-98/020. Available: <http://uwrl.usu.edu/people/faculty/dupont/epamonitoring.pdf>.
- Dupont, R.R., D.L. Sorensen, M. Kemblowski, M. Bertleson, D. McGinnis, I. Kamil and Y. Ma. 1998b. Project Summary: Monitoring and Assessment of In-Situ Biocontainment of Petroleum Contaminated Ground-Water Plumes. EPA/600/SR-98/020. Available: <http://www.p2pays.org/ref/07/06450.pdf>.
- Farhat, S.K., P.C. de Blanc, C.J. Newell, J.R. Gonzales and J. Perez. 2004. SourceDK Remediation Timeframe Decision Support System: User's Manual. Available: <http://www.gsi-net.com/en/software/free-software/sourcedk.html>.
- Faul, F., E.Erdfelder, A. Buchner and A.-G. Lang. 2009. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses, *Behavior Research Methods*, 41 (4), 1149-1160, doi:10.3758/BRM.41.4.1149. Both the journal article and the computer application are available: <http://www.psych.uni-duesseldorf.de/aap/projects/gpower/>.
- Hunkeler, D., R.U. Meckenstock, B. Sherwood Lollar, T.C. Schmidt and J.T. Wilson. 2008. A Guide for Assessing Biodegradation and Source Identification of Organic Ground Water Contaminants Using Compound Specific Isotope Analysis (CSIA). EPA 600/R-08/148 | December. Available: <http://www.epa.gov/ada/gw/mna.html>.
- Kampbell, D.H., C.B. Snyder, D.C. Downey and J.E. Hansen. 2001. Light nonaqueous-phase weathering at some JP-4 fuel release sites. *Journal of Hazardous Substance Research* 3(4), 1-7 . Available: http://www.engg.ksu.edu/hsrc/JHSR/v3_no4.pdf.
- Lu, X., D.H. Kampbell and J.T. Wilson. 2006. Evaluation of the Role of *Dehalococcoides* Organisms in the Natural Attenuation of Chlorinated Ethylenes in Ground Water. EPA/600/R-06/029. Available: <http://www.epa.gov/ada/gw/mna.html>.
- McHugh, T.E., L.M. Beckley, C.Y. Liu and C.J. Newell. 2011. Factors influencing variability in groundwater monitoring data sets. *Ground Water Monitoring & Remediation* 31(2): 92-101.
- Newell, C.J., H.S. Rifai, J.T. Wilson, J.A. Connor, J.A. Aziz and M.P. Suarez. 2002. Calculation and Use of First-Order Rate Constants for Monitored Natural Attenuation Studies. EPA/540/S-02/500. Available: <http://www.epa.gov/ada/gw/mna.html>.
- Newell, C.J., I. Cowie, T.M. McGuire, and W. McNab. 2006. Multi-year temporal changes in chlorinated solvent concentrations at 23 MNA sites. *Journal of Environmental Engineering* 132(6): 653–663.
- Peargin, T.R. 2001. Relative depletion rates of MTBE, benzene, and xylenes

from smear zone non-aqueous phase liquid. In: *Bioremediation of MTBE, Alcohols, and Ethers*. Proceedings of the Sixth International In Situ and On-Site Bioremediation Symposium. San Diego, California, June 4-7. Battelle Press 6(1):67-74.

- Pope, D., S. Acree, H. Levine, S. Mangion, J. van Ee, K. Hurt and B. Wilson. 2004. Performance Monitoring of MNA Remedies for VOCs in Ground Water. EPA/600/R-04/027. Available: <http://www.epa.gov/ada/gw/mna.html>.
- U.S. EPA. 1999. Use of Monitored Natural Attenuation at Superfund, RCRA Corrective Action, and Underground Storage Tank Sites. Office of Solid Waste and Emergency Response, Directive 9200.4-17P. Available: <http://www.cluin.org/download/reg/d9200417.pdf>.
- U.S. EPA. 2007. Treatment Technologies for Site Cleanup: Annual Status Report (Twelfth Edition) EPA-542-R-07-012. Available: http://www.clu-in.org/download/remed/asr/12/asr12_main_body.pdf.
- U.S. EPA. 2010. Superfund Remedy Report (Thirteenth Edition) EPA-542-R-10-004. Available: http://www.epa.gov/tio/download/remed/asr/13/srr_13th_maindocument.pdf.
- Wilson, J.T., P.M. Kaiser and C. Adair. 2005. Monitored Natural Attenuation of MTBE as a Risk Management Option at Leaking Underground Storage Tank Sites EPA-600/R-04/179. Available: <http://www.epa.gov/ada/gw/mna.html>.

Appendix A

First Phase Analysis, Using Linear Regression to Extract Rate Constants

This appendix provides further explanation about the statistical methods being used and detailed instructions on the use of a spreadsheet to extract the rate of attenuation from monitoring data.

The time required to reach a clean-up goal for a contaminant in ground water sampled by a monitoring well is directly related to the existing concentration of the contaminant in water, the value of the clean-up goal, and the rate constant for attenuation of that contaminant in water sampled by that well. Rate constants that are extracted from trends in concentration over time can be used to forecast the time that will be required for MNA to attain a particular clean up goal for a contaminant. The uncertainty in the estimate of the time required to reach the goal is directly related to the uncertainty in the estimate of the rate constant for attenuation in concentration over time. The OSWER Directive (U.S. EPA 1999) specifically requires: **Statistical confidence intervals should be estimated for calculated attenuation rate constants (including those based on methods such as historical trend data analysis, analysis of attenuation along a flow path in groundwater, and microcosm studies).** This appendix provides one approach to extract information from historical trend analysis in order to satisfy this requirement of the OSWER Directive.

The monitoring data are analyzed using linear regression. While specialized statistical software is not always available, spreadsheets are common computer applications. The simple statistical tests outlined here will be calculated using a spreadsheet rather than statistical software. The illustrations will be taken from calculations done with Microsoft Excel 2003; however, any spreadsheet can be used.

Section A.1 provides instructions to perform the linear regression on a particular data set. Section A.2 provides instructions on the use of Excel 2003 and links to computer applications available on the internet to determine whether a particular data set is appropriate for analysis using linear regression.

A.1. Use of the spreadsheet to calculate a linear regression

Open a spreadsheet to a blank sheet.

A.1.1. Express the sample dates as decimal years

Enter the dates when the samples were collected in a column, as was done in Column D of Figure A.1. Be sure that the cells are formatted as a date. To format as a date, left click with the mouse on the **D** column header to select the column, open **Format** from the menu bar, open the **Number** tab and select **Date** from the drop down menu. We intend to perform the regression on elapsed years in the data set, not elapsed days. To do that, we need to express the data as a decimal year. Copy the data from cells **D5** to **D31** and paste the data into cells **C5** through **C31**. Following the process above, format the cells in column **C** as a **Number** instead of a **Date**. Excel expresses the **Date** as the number of days since January 1, 1900. Insert a formula that converts the number of days in the **C** column into decimal years. This is done in cell **B5** in Figure A.1. Click on the cell to select it, then type the formula $=1900+C5/365.25$, which will appear in the *fx* box above the cells. Click on the check mark next to the *fx* box to accept the formula. Then left click on the square at the lower right of the cell with the formula [cell **B5** in Figure A.1], and drag it with the mouse to copy the formula into the other cells in the column. Column **B**

now contains the dates that the samples were collected expressed as decimal years.

(EXCEL USER NOTE: Microsoft Excel can calculate dates using two different date systems. One system (the one used here) sets January 1, 1900 as day 1. The other system sets January 1904 as day 1. Make sure you are not using the 1904 date system. Search on "date system" in Microsoft Office Help for instructions for your version of EXCEL.)

A.1.2. Calculate natural logarithms

Next, enter the reported values for the concentrations of TCE, as was done in cells **E5** through **E31** of Figure A.1. Make sure the cells are formatted as a **Number**. In cell **F5** insert a formula **=LN(E5)** to calculate the natural logarithm of the value in the adjacent cell. As above, copy the formula to the remaining cells in column **F**. Column **F** now contains the natural logarithm of the concentrations of TCE in column **E** in the spreadsheet.

	A	B	C	D	E	F	G
1							
2		Date	Date	Date	TCE	LN TCE	
3		Decimal Year	Number		µg/L		
4							
5		2000.9	36870	12/10/2000	1800	7.496	
6		2001.3	36985	4/4/2001	1200	7.090	
7		2001.5	37061	6/19/2001	2300	7.741	
8		2001.7	37157	9/23/2001	1600	7.378	
9		2001.9	37223	11/28/2001	880	6.780	
10		2002.3	37350	4/4/2002	1100	7.003	
11		2002.5	37431	6/24/2002	580	6.363	
12		2002.7	37523	9/24/2002	870	6.768	
13		2002.9	37594	12/4/2002	1400	7.244	
14		2003.2	37712	4/1/2003	1400	7.244	
15		2003.4	37777	6/5/2003	980	6.888	
16		2003.7	37894	9/30/2003	520	6.254	
17		2003.9	37960	12/5/2003	530	6.273	
18		2004.3	38084	4/7/2004	700	6.551	
19		2004.4	38146	6/8/2004	730	6.593	
20		2004.7	38252	9/22/2004	400	5.991	
21		2004.9	38322	12/1/2004	400	5.991	
22		2005.3	38454	4/12/2005	672	6.510	
23		2005.5	38516	6/13/2005	724	6.585	
24		2005.7	38604	9/9/2005	306	5.724	
25		2005.9	38698	12/12/2005	169	5.130	
26		2006.3	38812	4/5/2006	276	5.620	
27		2006.5	38889	6/21/2006	388	5.961	
28		2006.7	38967	9/7/2006	357	5.878	
29		2006.9	39057	12/6/2006	234	5.455	
30		2007.3	39180	4/8/2007	232	5.447	
31		2007.5	39252	6/19/2007	163	5.094	
32							

Figure A.1. A spreadsheet to calculate the decimal date of sampling and the natural logarithm of the concentration of TCE.

A.1.3. Run the regression

Next, click on **T**ools in the menu bar, and select the tab **D**ata **A**nalysis, then scroll down and select **R**egression from the menu. See Figure A.2. This will open an input screen as illustrated in Figure A.3.

(EXCEL USER NOTE: The instructions are for EXCEL 2003. If you are using another version of EXCEL, search on “data analysis” in Microsoft Office Help for instructions.

The Data Analysis ToolPak is not installed by default when EXCEL is installed. If the

Data Analysis tab is not shown under Tools, select Tools Add-Ins and check the two Analysis Tools options).

Input the data range of cell in column **F** (**LN TCE**) as the Input Y Range and the data range of cells in column **B** (**Date Decimal Year**) as the Input X Range. Input the desired confidence interval, and then click OK. The spreadsheet calculates the regression and presents output in a new tab in the spreadsheet (Figure A.4).

	A	B	C	D	E	F	G
1							
2		Date	Date	Date	TCE	LN TCE	
3		Decimal Year	Number		µg/L		
4							
5		2000.9	36870	12/10/2000	1800	7.496	
6		2001.3	36985	4/4/2001	1200	7.090	
7		2001.5	37061	6/19/2001	2300	7.741	
8		2001.7	37157	9/23/2001	1600	7.378	
9		2001.9	37223	11/28/2001	880	6.780	
10		2002.3	37350	4/4/2002	1100	7.003	
11		2002.5	37431	6/24/2002	580	6.363	
12		2002.7	37523	9/24/2002	870	6.768	
13		2003.9	37594	12/11/2002	1400	7.344	
14							
15							
16							
17							
18							
19							
20							
21							
22							
23							
24							
25		2005.9	38698	12/12/2005	169	5.130	
26		2006.3	38812	4/5/2006	276	5.620	
27		2006.5	38889	6/21/2006	388	5.961	
28		2006.7	38967	9/7/2006	367	5.878	
29		2006.9	39057	12/6/2006	234	5.455	
30		2007.3	39180	4/8/2007	232	5.447	
31		2007.5	39252	6/19/2007	163	5.094	

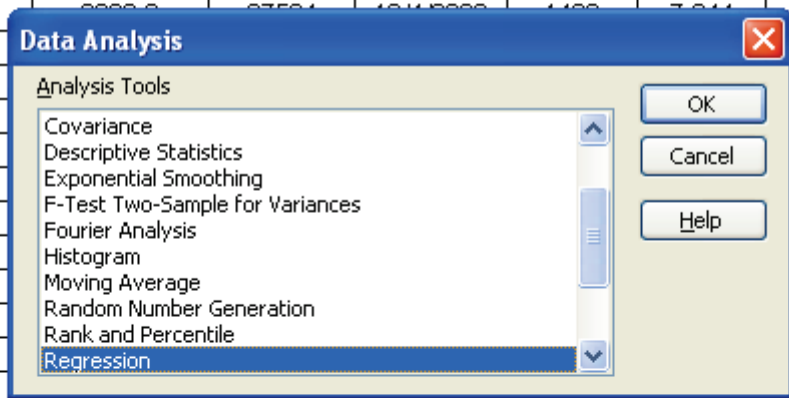


Figure A.2. The Regression dropdown menu under the Data Analysis tab in the Tools menu.

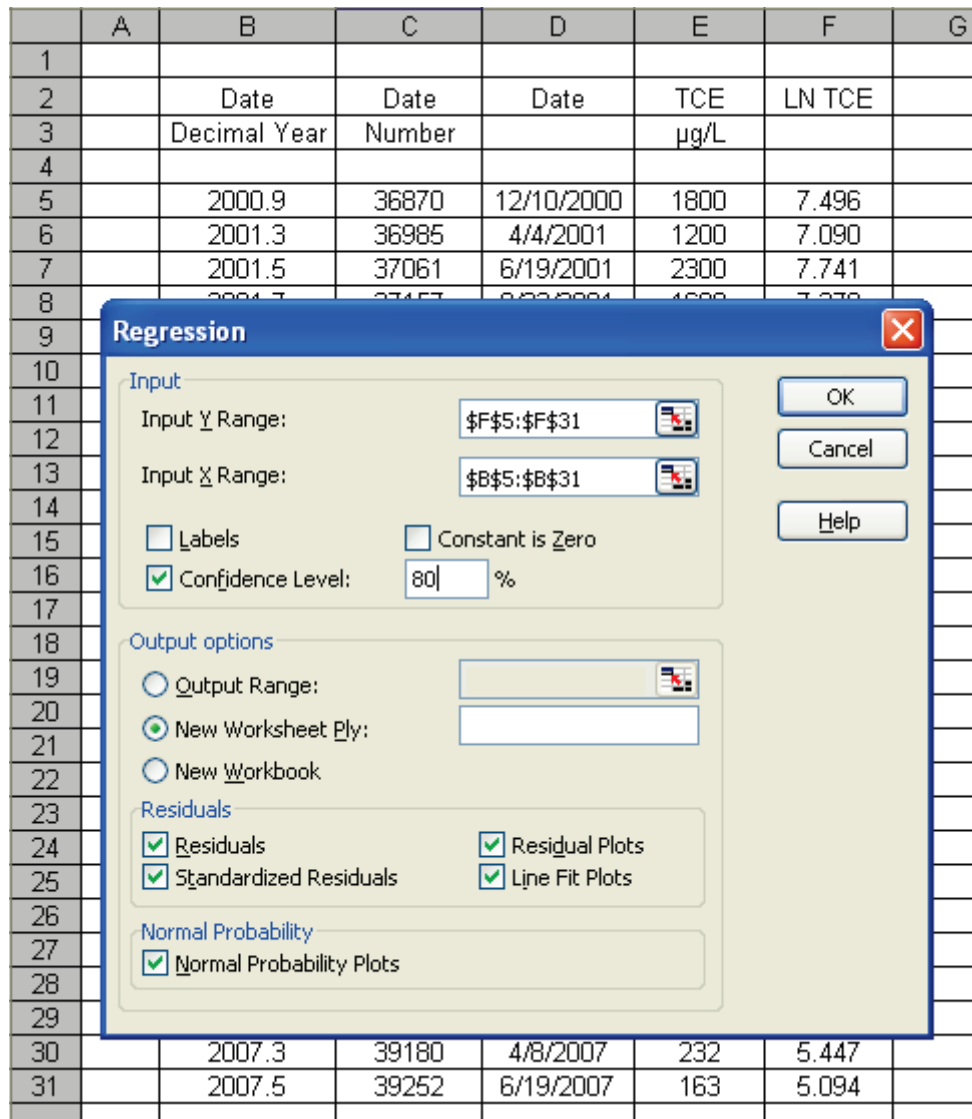


Figure A.3. The input menu for linear regression.

A.1.4. Examine the results

Examine Figure A.4. The slope of the regression is in cell **B18**, identified as the **Coefficient for X Variable 1**. The coefficient in the regression analysis is the first order rate constant for the rate of change. The rate of attenuation is the negative of the rate of change. For this particular data set, the first order rate of attenuation is +0.326 per year.

Examine Figure A.4. The values in cells **F18** and **G18** are the upper and lower 95% confidence intervals on the rate of change. The 95% confidence intervals are provided as a default in Excel 2003. Because we asked for an 80%

confidence interval, the values in cells **H18** and **I18** are the 80% confidence intervals. Because of the sign change, the “upper” confidence interval in cells is the slower confidence interval and the “lower” confidence interval in cells is the faster confidence interval.

The confidence intervals on the rates are calculated by Excel using the Student’s *t* distribution. The calculations use a two-tailed value for the *t* statistic. That means that one half the error is in rates that are faster than the “faster” confidence interval, and one half is in rates that are slower than the “slower” confidence interval. A rate faster than is expected is a desirable

	A	B	F	G	H	I
1	SUMMARY OUTPUT					
2						
3	<i>Regression Statistics</i>					
4	Multiple R	0.886				
5	R Square	0.784				
6	Adjusted R Square	0.776				
7	Standard Error	0.346				
8	Observations	27.000				
9						
10	<i>ANOVA</i>					
11		<i>df</i>	<i>Significance F</i>			
12	Regression	1.000	0.00000000083			
13	Residual	25.000				
14	Total	26.000				
15						
16		<i>Coefficients</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 80.0%</i>	<i>Upper 80.0%</i>
17	Intercept	659.700	518.563	800.837	569.493	749.907
18	X Variable 1	-0.326	-0.396	-0.256	-0.371	-0.281

Rate of attenuation
= 0.326 per year

Slower 90%
confidence interval
on the rate of
attenuation
= 0.281 per year

Figure A.4. The summary output of the linear regression.

outcome, and there is no need to use statistics to “protect” ourselves from the possibility of faster rates of attenuation. The best course is to put all the uncertainty on the side with the slower rate. Statisticians refer to this as a one-tailed test because all the uncertainty is in one tail of the frequency distribution. The term 80.0% confidence in the Excel spreadsheet means we are willing to accept a 20% chance of error, 10% of the chance in the faster tail and 10% in the slower tail. As a result, the two-tailed Upper 80.0% confidence interval in cell **I18** of Figure A.4 is also the slower one tailed 90% confidence interval on the rate. As revealed by the regression analysis in Figure A.4, the rate constant for attenuation of TCE concentrations over time is 0.326 per year (cell **B18**), and that rate constant is at least 0.281 per year at 90% confidence (cell **I18**). Remember that the rate of attenuation is the negative of the rate of change as calculated in EXCEL.

A.2. Statistical background on linear regression

Data analysis using linear regression makes three major assumptions:

- 1) there is a linear relationship between the dependent variable (concentration) and independent variable (time or date of sampling),
- 2) the variance in the data is constant over time, and
- 3) the residual errors follow a normal probability distribution.

The regression in Section A.1 was conducted on a natural logarithm transformation of the original concentration data. This section explains why the transformation was appropriate for the example data set, and provides an approach to determine if the natural logarithm transformation is appropriate for other data sets.

A.2.1. Linear relationship between variables

If the relationship is not linear, this is usually most evident in a plot of the observed versus predicted values. The data points should be symmetrically distributed around the regression line. Look carefully for evidence that the data follow a “bowed” pattern, which would indicate that the data do not follow a linear model. This approach is illustrated in Figure A.5.

Panels A and B of the figure were created as charts in Excel 2003. The spreadsheet provides a convenient option of fitting linear, exponential, logarithmic and polynomial trend lines to the data in the chart, and extracting an equation and values of r^2 for the trend line. The Coefficient of Determination (r^2) is an estimate of the goodness of fit of the data to the equation. The value of the coefficient is the fraction of the total variation in the data that is explained by the equation. If r^2 were 0.800, then 80% of the variation in the dependant variable is explained by the equation that is fit to the data, and 20% remains unexplained by the equation.

Panel A of Figure A.5 plots the data on an arithmetic scale, and fits a linear regression line. Notice that the distribution of the example data seems to be bowed. Most of the data from the very earliest dates are above the line, and in one case the datum is far above the line, most of the data from the intermediate dates are below the line, and most of the data from the latest dates are again above the line. The data are better fit with an exponential trend line than a linear trend line ($r^2 = 0.7843$ for the exponential trend line in Panel B of Figure A.5 compared to $r^2 = 0.6848$ for the linear trend line in Panel A of Figure A.5). The value of r^2 for a logarithmic trend line was 0.6849, and the value for a polynomial trend line varied from $r^2 = 0.7199$ for a second degree polynomial to $r^2 = 0.7266$ for a six degree polynomial (data not presented). Of the available trend lines in Excel 2003, the exponential trend line has the highest r^2 and provides the best fit to the example data set.

Figure A.6 is a linear trend line on the natural logarithm of the concentrations of TCE. Notice that the data now follow the linear trend line without any apparent curvature. Taking a logarithm reduces an exponent in the dependant variable to a factor, making the relationship between the variables a linear function of the independent variable. A logarithmic transformation of the concentration data will provide a better fit to the assumptions of linear regression. The value of r^2 for the linear trend of the natural logarithm of TCE on date of sampling is the same as the exponential trend of concentration on date. The natural logarithm transformation of the data provides a better fit of the assumption of a linear relationship.

A.2.2. Uniform variance in data

The vertical distance between each data point and the regression line is called the residual variation or residual for that data point. If the variance in the data is constant over time, the residuals will be uniformly distributed along the regression line. Notice in Panel A of Figure A.5 that the distances are greater at earlier dates and smaller at later dates, while in Panel B the distances are variable but uniformly spread along the line. Obviously, the natural logarithms of the concentrations provide a better fit to the assumption of uniform variance in the data with date of sampling.

A.2.3. Normal distribution of residuals

The third criterion, that the residuals from the regression follow a normal probability distribution, is not readily apparent from a simple inspection of a plot of the data. Excel 2003 gives the user an option to view a table with the residuals from the regression. The menu depicted in Figure A.3 provides an option to see a table of the **Residuals**, and the **Standard Residuals** from the regression. Figure A.7 presents the RESIDUAL OUTPUT report provided in Excel 2003 when these options are selected. As discussed above, the numbers in the column labeled *Residuals* are simply the difference between the value of the dependant variable and the predicted value of the

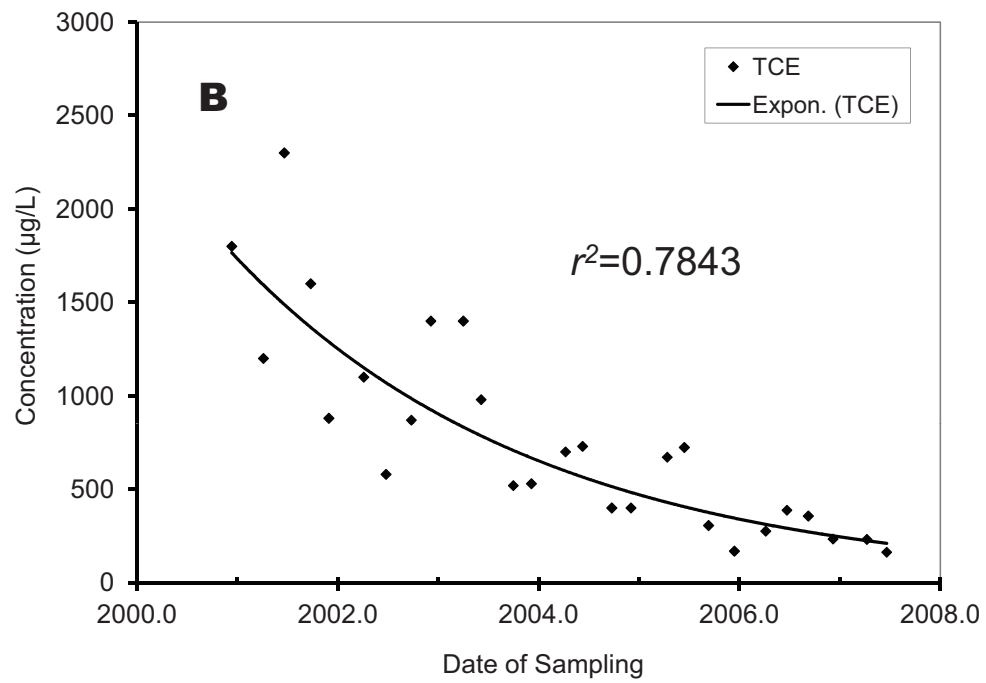
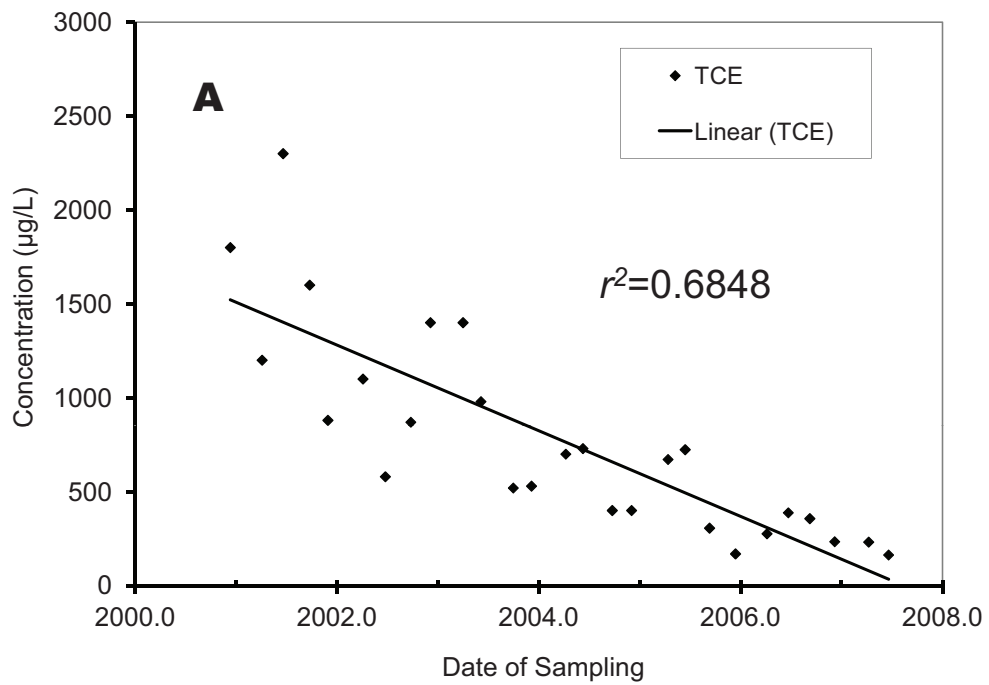


Figure A.5. Fit of a linear trend and an exponential trend to concentration of TCE.

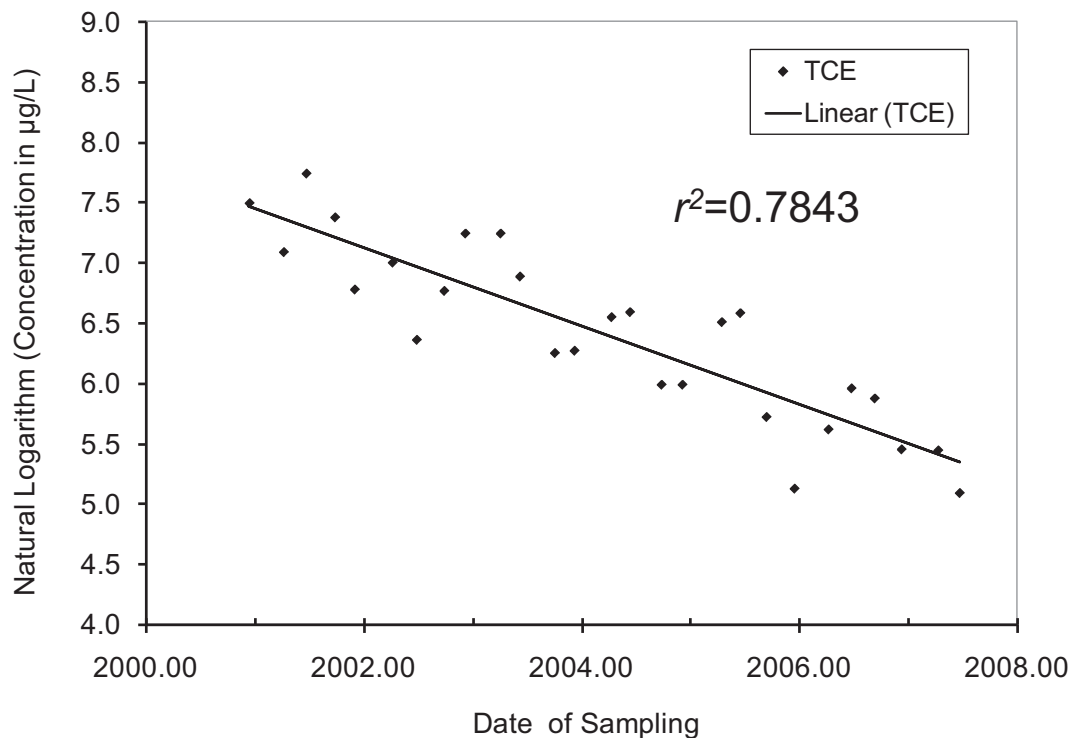


Figure A.6. Fit of a linear trend to the natural logarithm of concentration of TCE.

dependant variable (*Predicted Y*) in the report as depicted in the figure.

The assumption of a normal distribution of residuals is evaluated by comparing the distribution of the residuals to the normal probability distribution. The first step is to examine a normal probability plot or Q-Q plot to see if there are obvious differences between the distribution of the residuals of regression and the normal distribution. The second step is a goodness-of-fit test between the residuals and the normal distribution. Two approaches to make these comparisons are discussed in Sections A.3. and A.4. below. Examine both sections to identify the approach that is most appropriate for you.

A.3. Using Excel to generate a Q-Q plot and using the One-Sample Kolmogorov-Smirnov Goodness-of-Fit Test

The standard probability distribution has a mean of zero and a standard deviation of one. The regression calculates an equation such that the

mean of the *Residuals* is zero. In the column labeled *Standard Residuals*, the *Residuals* have been scaled to make their standard deviation also equal to one. The numbers labeled PROBABILITY OUTPUT are not needed for this analysis. Delete this information to create space in the spreadsheet for a calculation of the values of the standard normal probability distribution that correspond to the *Standard Residuals*.

Select the values under *Standard Residuals*, then open the **Data** drop down menu from the menu bar, select **Sort**, select **Expand the Selection**, and then the **Sort** button. Indicate that you intend to sort with a header, and sort ascending. The *Observation* number, *Predicted Y*, *Residuals*, and *Standard Residuals* should all sort together.

Label the column immediately to the right of *Standard Residuals* as *Rank*, and insert numbers to order the *Standard Residuals* from least to greatest, as illustrated in Figure A.8. Label the next column to the right as *Quantile*. In the first cell under *Quantile* insert a formula for the

	A	B	C	D	E	F	G
22	RESIDUAL OUTPUT					PROBABILITY OUTPUT	
23							
24	<i>Observation</i>	<i>Predicted Y</i>	<i>Residuals</i>	<i>Standard Residuals</i>		<i>Percentile</i>	<i>Y</i>
25	1	7.476268292	0.01927365	0.056823429		1.851851852	5.093750201
26	2	7.373639551	-0.28356272	-0.836012096		5.555555556	5.129898715
27	3	7.305815339	0.43484906	1.282041175		9.259259259	5.446737372
28	4	7.220142651	0.15761626	0.464691199		12.96296296	5.455321115
29	5	7.161242677	-0.38132077	-1.124226702		16.66666667	5.620400866
30	6	7.04790485	-0.04483939	-0.132197469		20.37037037	5.723585102
31	7	6.975618519	-0.61259042	-1.806066066		24.07407407	5.877735782
32	8	6.893515526	-0.12502231	-0.368596297		27.77777778	5.96100534
33	9	6.830153433	0.41407408	1.220791463		31.48148148	5.991464547
34	10	6.72484742	0.5193801	1.531259292		35.18518519	5.991464547
35	11	6.666839871	0.2207127	0.650714914		38.88888889	6.253828812
36	12	6.562426282	-0.30859747	-0.909820664		42.59259259	6.272877007
37	13	6.503526308	-0.2306493	-0.680010438		46.2962963	6.363028104
38	14	6.392865752	0.15821458	0.466455207		50	6.510258341
39	15	6.337535474	0.25550906	0.753303074		53.7037037	6.551080335
40	16	6.242938547	-0.251474	-0.741406732		57.40740741	6.584791392
41	17	6.180468879	-0.18900433	-0.557230901		61.11111111	6.593044534
42	18	6.062668932	0.44758941	1.319602825		64.81481481	6.768493212
43	19	6.007338654	0.57745274	1.702471618		68.51851852	6.779921907
44	20	5.928805356	-0.20522025	-0.605039399		72.22222222	6.887552572
45	21	5.844917515	-0.7150188	-2.108049945		75.92592593	7.003065459
46	22	5.743181198	-0.12278033	-0.361986387		79.62962963	7.090076836
47	23	5.674464562	0.28654078	0.844792151		83.33333333	7.244227516
48	24	5.604855503	0.27288028	0.804517668		87.03703704	7.244227516
49	25	5.524537357	-0.06921624	-0.204066375		90.74074074	7.377758908
50	26	5.414769225	0.03196815	0.094249899		94.44444444	7.495541944
51	27	5.350514709	-0.25676451	-0.75700444		98.14814815	7.740664402

Figure A.7. Report in Excel 2003 of the Residuals and Standard Residuals from the Linear Regression of the example data set.

center of the interval represented by rank. The formula is the rank minus 0.5, then that difference divided by the number of observations. Drag the formula to paste into all the cells with an observation.

Then label the column immediately to the right of *Quantile* as *z-Score*. Insert a formula that returns the value of the z-distribution that corresponds to the appropriate quantile. For Excel 2003, the formula is =NORMSINV(column row). Drag the formula to paste into all the cells with an observation.

The distribution of the *Standard Residuals* will be compared to the standard normal distribution (*z-Score*), to determine if the distributions

are different. The first comparison is graphical. Create a chart with two series of data. The first series plots values under *z-Score* as the x-axis and the *Standard Residuals* as the y-axis. The second series plots *z-Score* as both the x-axis and y-axis. The resulting chart compares the distribution of the *Standard Residuals* to the normal probability distribution. The normal probability distribution will lie along a straight line. If the *Standard Residuals* are normally distributed, they will lie along the same line. Jumps or breaks between adjacent data points may indicate that the data are sampled from two different populations with different properties. A sample that is well separated from the line may be an “outlier” (Singh et al. 2010). The

$=(E25-0.5)/27$ $=\text{NORMSINV}(F25)$

	A	B	C	D	E	F	G
21							
22	RESIDUAL OUTPUT						
23							
24	<i>Observation</i>	<i>Predicted Y</i>	<i>Residuals</i>	<i>Standard Residuals</i>	<i>Rank</i>	<i>Quantile</i>	<i>z-Score</i>
25	21	5.845	-0.715	-2.108	1	0.019	-2.085
26	7	6.976	-0.613	-1.806	2	0.056	-1.593
27	5	7.161	-0.381	-1.124	3	0.093	-1.325
28	12	6.562	-0.309	-0.910	4	0.130	-1.128
29	2	7.374	-0.284	-0.836	5	0.167	-0.967
30	27	5.351	-0.257	-0.757	6	0.204	-0.828
31	16	6.243	-0.251	-0.741	7	0.241	-0.704
32	13	6.504	-0.231	-0.680	8	0.278	-0.589
33	20	5.929	-0.205	-0.605	9	0.315	-0.482
34	17	6.180	-0.189	-0.557	10	0.352	-0.380
35	8	6.894	-0.125	-0.369	11	0.389	-0.282
36	22	5.743	-0.123	-0.362	12	0.426	-0.187
37	25	5.525	-0.069	-0.204	13	0.463	-0.093
38	6	7.048	-0.045	-0.132	14	0.500	0.000
39	1	7.476	0.019	0.057	15	0.537	0.093
40	26	5.415	0.032	0.094	16	0.574	0.187
41	4	7.220	0.158	0.465	17	0.611	0.282
42	14	6.393	0.158	0.466	18	0.648	0.380
43	11	6.667	0.221	0.651	19	0.685	0.482
44	15	6.338	0.256	0.753	20	0.722	0.589
45	24	5.605	0.273	0.805	21	0.759	0.704
46	23	5.674	0.287	0.845	22	0.796	0.828
47	9	6.830	0.414	1.221	23	0.833	0.967
48	3	7.306	0.435	1.282	24	0.870	1.128
49	18	6.063	0.448	1.320	25	0.907	1.325
50	10	6.725	0.519	1.531	26	0.944	1.593
51	19	6.007	0.577	1.702	27	0.981	2.085

Figure A.8. Calculation of the values of the standard normal distribution (z-distribution) corresponding to the rank and quantile of the standard residuals.

data in the outlier could be real, or it could be an error in labeling the sample in the field, an error in reading the label in the laboratory, an error in dilution of the sample in the laboratory, or an error in entering the data in the data file.

The plot for the example data is presented in Figure A.9. There is little evidence that the *Standard Residuals* are not normally distributed.

Excel 2003 has a menu option for Normal Probability Plots. Excel does not provide a normal probability plot or Q-Q plot as the terms

are usually understood. Instead Excel provides a plot of the original data against the quantile expressed as percent of the number of samples. Construct your own normal probability plot as described above, and ignore the “normal probability” plot provided by Excel.

Interpreting a normal probability plot requires a certain amount of judgment, which in turn is based on experience. If you want a more objective means to compare the distributions, there are a number of statistical tests that can be used to compare the distribution of *Standard*

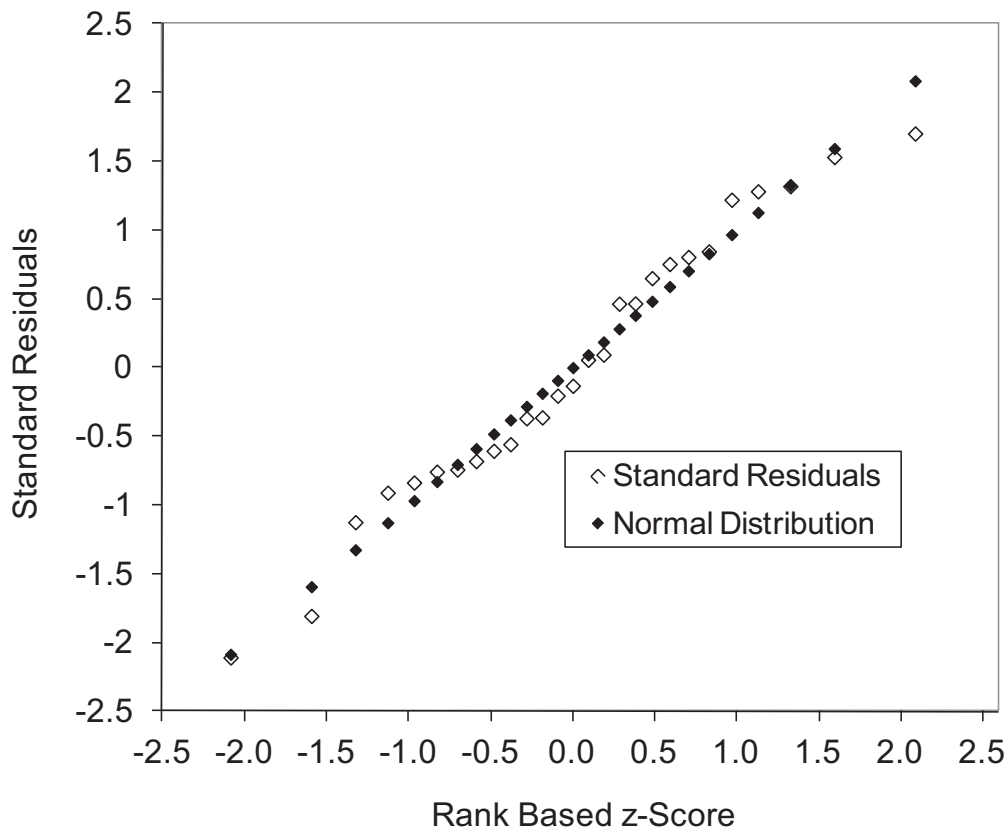


Figure A.9. Normal Probability Plot (also called a Q-Q plot) comparing the distribution of the residuals from the regression to the normal probability distribution.

Residuals to the normal probability distribution. One useful statistic is the One-Sample Kolmogorov-Smirnov Goodness-of-Fit Test. The test is also called the Kolmogorov-Smirnov D test or Kolmogorov-Smirnov Z test. The Kolmogorov-Smirnov Test provides an estimate of the probability that a given distribution is not significantly different from a second distribution. In this case, the first distribution will be the *Standard Residuals* from the regression and the second will be the normal distribution as described by the *z-Scores*.

Applications to perform the Kolmogorov-Smirnov Test are readily available on the internet. One that is particularly useful can be found at <http://www.physics.csbsju.edu/stats/KS-test.html>. To use the application, open the KS-test Data Entry form provided as a link from the web page and copy the values for *Standard Residuals* from the Excel spreadsheet and paste

into Dataset 1 and copy values from *z-Score* from Excel and paste into Dataset 2, then calculate (see Figure A.10). The application returns with the following statement regarding distribution 1, which is the *Standard Residual* after regression: *KS finds the data is consistent with a normal distribution: P = 0.93 where the normal distribution has mean= -2.6199E-02 and sdev= 1.106.* See Figure A.11.

For any value of *P* above 0.50, the Kolmogorov-Smirnov Test finds that it is more likely that the distribution of *Standard Residuals* is not significantly different from a normal distribution. This leads to the following suggested decision criterion. **For any value of *P* above 0.50, it is most likely that the *Standard Residuals* follow a normal distribution, and the approach and procedures in Appendix A are appropriate. If the value of *P* is less than 0.50, it is most likely that the**

KS-test Data Entry

Use the below form to enter your data for a Kolmogorov-Smirnov test. The KS-test seeks differences between your two datasets; it is non-parametric and distribution free. Reject the null hypothesis of no difference between your datasets if P is "small". In addition this page reports if your datasets seem to have normal or lognormal distribution. This may allow you to use other tests like the t -test.

For each dataset, enter your data into the given box separating each datum from its neighbor with tabs, commas, or spaces. Very commonly you will already have the data in your computer in some format. You should be able to just copy and paste that data into the appropriate area. This KS-test form is designed to handle datasets with between 10 and 1024 items in each dataset.

Dataset 1: ← Standard Residuals

```
-2.108049945  
-1.806066066  
-1.124226702  
-0.909820564  
-0.836012096  
-0.75700444
```

Dataset 2: ← Z-Scores

```
-2.085355566  
-1.593218318  
-1.324957589  
-1.128143545  
-0.967421566  
-0.828464549
```

Figure A.10. Data entry form in an application to perform the Kolmogorov-Smirnov Test.

Standard Residuals do not follow a normal distribution, and the approach and procedures in Appendix A should not be used with that data set to extract a first order rate constant for natural attenuation.

A.4. Using ProUCL for a Q-Q plot and Goodness-of-Fit testing

U.S. EPA provides free software that can easily construct a Q-Q plot and test for goodness-of-fit. Download and install ProUCL Version 4.00.05 (Singh et al., 2010). Copy the column heading and data for residuals from regression from column C in the Excel spreadsheet as

depicted in Figure A.8 and paste the data into a new Excel workbook. Format the data as a number and the column label as text. Save the workbook as an Excel 97-2003 workbook (*.xls) in the *data* file in the directory containing ProUCL. Open the **File** menu in ProUCL, select the **Load Excel Data** option, select the workbook and open the data into ProUCL. Open the menu **Graphs** and select **Multi Q-Q** to open the menu **Select Variables**. Select **Residuals** and click OK. ProUCL generates a Q-Q plot (Figure A.12, compare to Figure A.9).

ProUCL also will do goodness-of-fit testing. Open the menu **Goodness-of-Fit**, to open the

KS Test: Results

Kolmogorov-Smirnov Comparison of Two Data Sets

The results of a Kolmogorov-Smirnov test performed at 17:08 on 6-JUL-2011

The maximum difference between the cumulative distributions, D , is: 0.0741 with a corresponding P of: 1.000

Data Set 1:

27 data points were entered

Mean = 1.1111E-10

95% confidence interval for actual Mean: -0.3956 thru 0.3956

Standard Deviation = 1.000

High = 1.70 Low = -2.11

Third Quartile = 0.805 First Quartile = -0.741

Median = -0.1322

Average Absolute Deviation from Median = 0.824

KS finds the data is consistent with a normal distribution: $P=0.93$ where the normal distribution has mean= -2.6199E-02 and sdev= 1.106

Figure A.11. Results returned from the Kolmogorov-Smirnov test.

menu **Select Variables**. Select **Residuals**. Open **Options** and select the desired level of confidence. Click OK to **Select Confidence Level** and OK to **Select Variables** and ProUCL produces values for the Shapiro Wilk (SW) Test Statistic and the Lilliefors Test Statistic (Figure A.13). According to Singh et al. (2010) *ProUCL 4.0 provides S-W test only for samples of sizes up to 50. Lilliefors test (along with graphical Q-Q plot) seems to perform fairly well for samples of size 50 and higher.* Both tests are restricted in ProUCL to confidence levels of 90%, or 95% or 99%. If the data appear to be normal at $\alpha = 0.10$, there is no reason to reject the Approach in this Report. However, the limited menu for statistical confidence provided by ProUCL 4.0 does not allow ProUCL to evaluate the distribution of residuals against the decision criterion of a probability of 50%.

References:

Singh, A., R. Maichle, N. Armbya, A.K. Singh and S.E. Lee. 2010. ProUCL Version 4.00.05 User Guide (Draft). EPA/600/R-07/038. Available <http://www.epa.gov/esd/tsc/software.htm>, updates available at <http://www.epa.gov/osp/hstl/tsc/software.htm>.

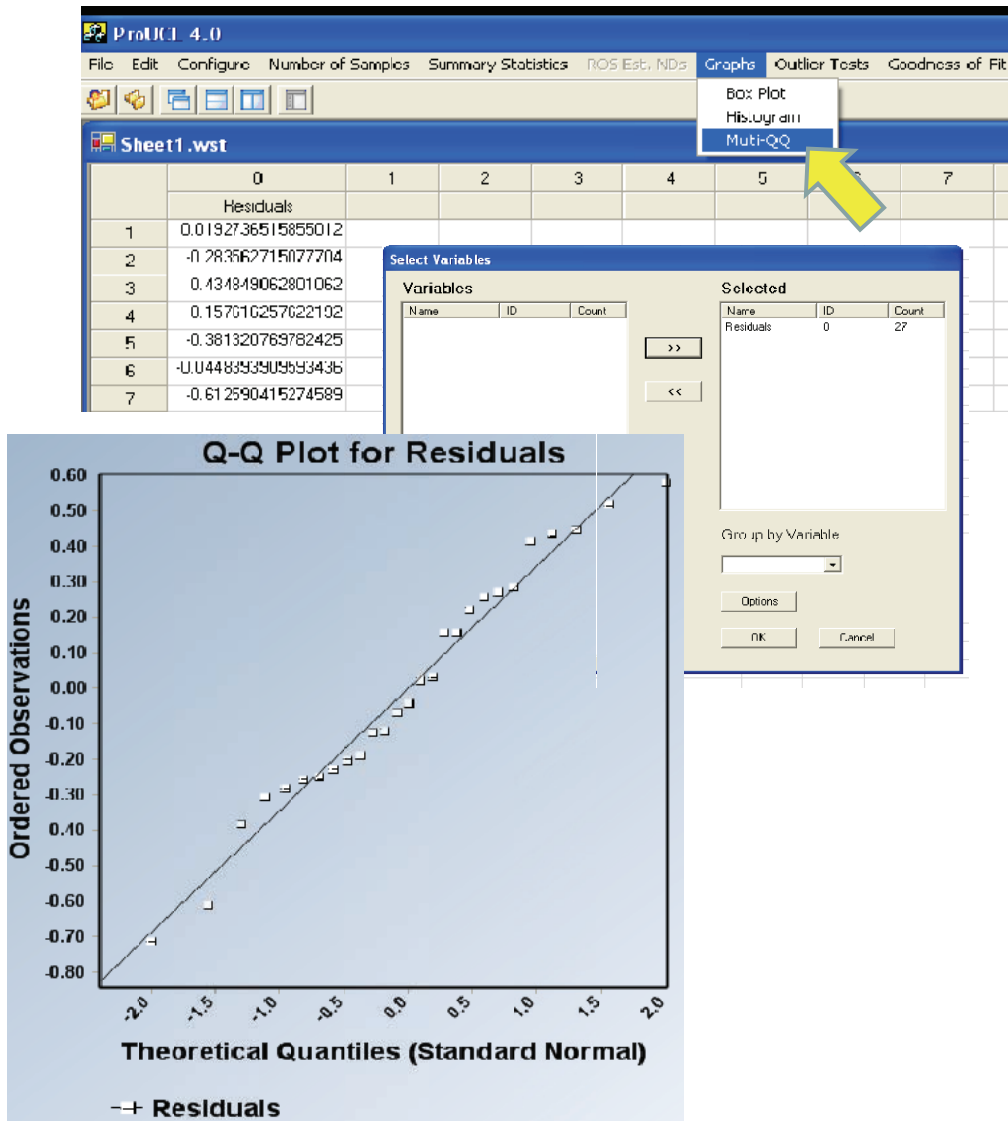


Figure A.12. Using ProUCL to generate a Q-Q plot.

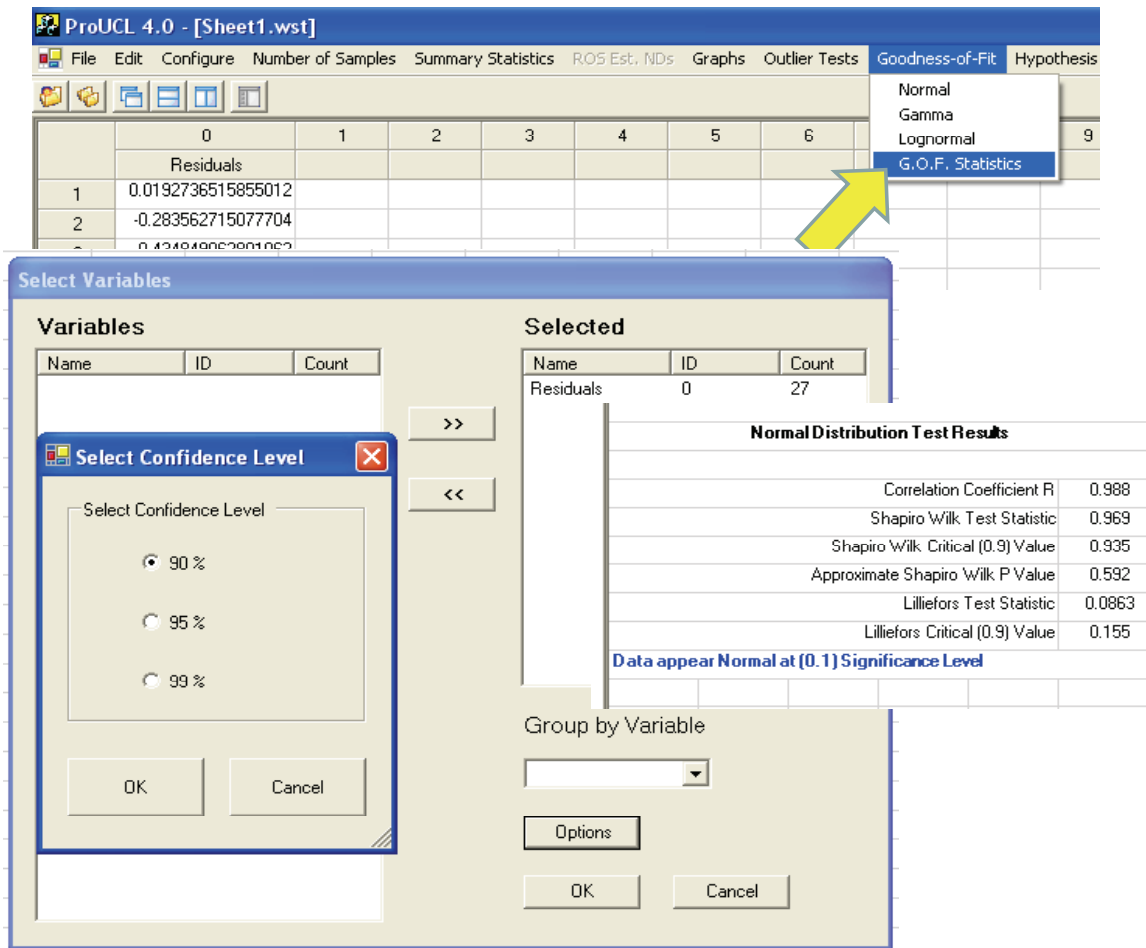


Figure A.13. Using ProUCL to evaluate goodness-of-fit to a normal distribution.

Appendix B

Second Phase Analysis

B.1. Statistical Approach

The purpose of the second phase analysis is to compare conditions at the beginning of a review cycle to conditions at the end of the review cycle. Instead of a linear regression of all the concentration data collected in the review cycle over time, we will only compare data collected in the first year of the cycle to data collected in the final year of the cycle. To have statistical control on the extent of attenuation from the first to the last year, it is necessary to have multiple samples of the concentration of the contaminant in the water from the well in both the first year and in the final year. The samples in the first year will be used to calculate a population of expected concentrations in the last year of the review cycle, based on

- 1) a concentration-based long term goal for clean up,
- 2) a date by which the goal is to be obtained, and
- 3) the assumption that attenuation will follow first order kinetics.

These expectations will be taken as interim goals.

The approach assumes that each year of the monitoring record can be treated as a single entity, and the effects of attenuation within each year can be ignored. McHugh et al. (2011) evaluated three large monitoring records, and found that there was a characteristic time-independent variability associated with subsequent samples from the same well, and that long term trend in attenuation in concentration could not be distinguished from the time-independent variability unless the samples were separate in time by more than 320 to 400 days. As an acceptable approximation, the effects of attenuation within a monitoring year will be ignored.

All calculations are performed on the natural logarithm of the concentrations in the samples. A geometric mean will be calculated for the interim goals and for the samples in the final year of the review cycle. A Student's t test for the difference of means will be used to compare the mean for the interim goals to the geometric mean of samples in the final year of the review cycle. This test provides a confidence interval on the difference between the mean of the interim goals and the mean of the final year of data. If the mean of samples in the final year is greater than the mean of the interim goals at some predetermined level of confidence, that fact will indicate that attenuation in the review cycle is not adequate to meet the long term goal. The Student's t test for the difference of means performed on the natural logarithm of the concentration data is equivalent to comparing the ratio of the means to determine if the attenuation ($C_{\text{final}}/C_{\text{initial}}$) is adequate to attain the goals.

B.1.1. Statistical background: Use of the Student's t -statistic

The Student's t -statistic is simply the ratio of "signal" to "noise" in a data set. In this application, the calculated value of the t -statistic is the difference between the means divided by the standard deviation of the difference. The difference in the means is the signal and the standard deviation is a measure of the noise. A difference between means can be tested to see if it is statistically different from zero by comparing the calculated value of the t -statistic to the critical value of t . The critical value of t is the maximum value of t that can be expected from a theoretical distribution of t where:

- 1) the variation in the measured values is due only to normal random variation, and

2) the user is willing to accept a predetermined chance that the test will fail and conclude that the difference in the means is different from zero when in fact it is not different.

If the predetermined chance that the test will fail is α , then the level of confidence in the test is $(1 - \alpha)$. As an example, a value of α of 0.05 would correspond to a confidence level of 95%; a value of α of 0.20 would correspond to a confidence level of 80%. If the calculated t is greater than the critical value of t , the difference in the means is statistically different from zero at that predetermined level of confidence.

Statistics text books provide tables of the critical value of t calculated for predetermined values of α such as 0.01, 0.05, 0.10, or 0.15. Spreadsheets have macros that can calculate the critical value of t for any arbitrary value of α .

B.1.2. Theoretical basis of the statistical comparison of means

In this application, the signal is the difference in the means of the initial year and final year of the review cycle. Student's t is a calculated statistic that will be compared to the critical value of t which depends on the acceptable probability of error and the number of degrees of freedom in the comparison ($t_{\alpha, d.f.}$). For this case, t is defined in Equation B.1, where \bar{X}_1 is the mean of the samples collected in the initial year, \bar{X}_2 is the mean of the samples collected in the final year, μ_1 is the true but unknown mean concentration in the initial year, μ_2 is the true but unknown mean in the final year, and $s_{(\bar{X}_1 - \bar{X}_2)}$ is the standard deviation of the difference of the means.

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{s_{(\bar{X}_1 - \bar{X}_2)}} \quad \text{B.1}$$

If we postulate that the real difference in the means in the initial and final year is zero, where $(\mu_1 - \mu_2) = 0$, then

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{s_{(\bar{X}_1 - \bar{X}_2)}} \quad \text{and}$$

$$(\bar{X}_1 - \bar{X}_2) - 0 = s_{(\bar{X}_1 - \bar{X}_2)} \times t_{\alpha, d.f.}$$

The difference of means is significantly greater than zero whenever Equation B.2 pertains.

$$(\bar{X}_1 - \bar{X}_2) > s_{(\bar{X}_1 - \bar{X}_2)} \times t_{\alpha, d.f.} \quad \text{B.2}$$

In equation B.2, $t_{\alpha, d.f.}$ is the critical value of the t -statistic for a particular value of α and the appropriate number of degrees of freedom in the comparison. When there are the same number of samples in each mean and the variance of the two means are the same, the degrees of freedom are simply $2(n-1)$ where n is the number of samples in each mean. When there are a different number of samples in each mean or the variance of the means is different, the calculations are more complex.

The demonstration spreadsheet (EvalMNA.xls) which accompanies this appendix allows for the possibility that there are a different number of samples in each mean, and that the variance of the means is different. When the number of samples in the means are different, or the variances are different, the standard deviation of the difference between the means is calculated following equation B.3, where s_{X_1} is the sample standard deviation of the first mean, s_{X_2} is the sample standard deviation of the second mean, n_1 is the number of measurements in the first mean, and n_2 is the number of measurements in the second mean:

$$s_{(\bar{X}_1 - \bar{X}_2)} = \left[\frac{s_{X_1}^2}{n_1} + \frac{s_{X_2}^2}{n_2} \right]^{1/2} \quad \text{B.3}$$

The number of degrees of freedom in $s_{(\bar{X}_1 - \bar{X}_2)}$ can be approximated by the Welch-Satterthwaite equation, Equation B.4 below.

$$d.f. = \frac{(s_{X_1}^2 / n_1 + s_{X_2}^2 / n_2)^2}{(s_{X_1}^2 / n_1)^2 / (n_1 - 1) + (s_{X_2}^2 / n_2)^2 / (n_2 - 1)} \quad \text{B.4}$$

B.1.3. Use of a spreadsheet to calculate the difference of means

The following provides a detailed step by step explanation of the application of the t -statistic for the difference of means to determine whether the difference in concentrations between the initial and final years of the review cycle are consistent with meeting the cleanup goals. If you have access to the internet, download the spreadsheet from <http://www.epa.gov/nrmrl/gwerd/csmos/models/EvalMNA.xls>. Open the tab *MNA Evaluation* and enter data from the evaluation. If the spreadsheet is not available,

reconstruct the spreadsheet from a blank spreadsheet, following instructions provided below.

To set up the spreadsheet from a blank spreadsheet, enter equations in the appropriate cells. Figure B.1 shows the equations to be entered in cells **D13** through **D16**, cells **D23** through **D26**, cell **E15**, cell **E25**, cell **F15** and cell **F25**.

The equation in **D13** should be copied into cells **D14** through **D16** and **D23** through **D26**. Click on cell **D13** to select it, and then type in the formula $=LN(C13)$. The

	A	B	C	D	E	F
1						
2		Comparison of Initial Year of Review Cycle t				
3						
4		Date	Concentration	LN Conc.	Mean	Geometric
5			µg/L		LN Conc.	Mean of
6						Conc. (µg/L)
7						
8		=LN(C13)	=AVERAGE(D13:D16)	=EXP(E15)		
9						
10		Initial Year	Initial Year	Initial Year	Initial Year	Initial Year
11		$n_1=?$				
12		4				
13		4/4/2001	1200	7.090		= C_0
14		6/19/2001	2300	7.741		
15		9/23/2001	1600	7.378	7.247	1404
16		11/28/2001	880	6.780		
17						
18			=LN(C23)			
19						
20		Final Year	Final Year	Final Year	Final Year	Final Year
21		$n_2=?$				
22		4				
23		4/5/2006	276	5.620		= C_i
24		6/21/2006	388	5.961		
25		9/7/2006	357	5.878	5.729	308
26		12/6/2006	234	5.455		
27						
28			=AVERAGE(D23:D26)	=EXP(E25)		
29						
30						

Figure B.1. Populating the spreadsheet *Evaluation of MNA* to calculate the mean concentrations of contaminant in the first and in the final year of the review cycle. The data entry cells are formatted in red. The formulas in the cells identified with arrows are provided in the blue text boxes.

formula will appear in the fx box above the cells. Click on the check mark to accept the formula. Then left click on the square at the lower right of the cell with the formula [cell D13 in Figure B.1], and drag it with the mouse to copy the formula into cells D14 through D16. Copy the formula in cell D13 and paste into cell D23. Then drag the formula into cells D24 through D26.

Continue to create the spreadsheet by entering formulas in cells as described in Figures B.2 through B.5. Formulas will

have to be copied and dragged into cells as described above in columns M, N, O, S, T, Z, AB, AC and AD. The use of the \$ in the formulas fixes the row number as a formula is copied into other cells. Site specific data must be entered into cells P13, Q13, and R13. The user must enter data in Column K to specify the acceptable probability of error in the comparison. The values in Figure B.3 are for illustration only. The font used in Excel was Arial, and as a result it is important to distinguish the letter I used in the formulas from the number 1. If the spreadsheet

	D	E	F	G	H	I	J
1							
2	Comparison of Initial Year of Review Cycle to Final Year of Review Cycle						
3							
4	LN Conc.	Mean	Geometric	Stan Dev	Difference	Stan Dev	Attenuation
5		LN Conc.	Mean of	LN Conc.	of Means	Difference	Factor
6			Conc. (µg/L)		(LN Conc.)	of Means	C/Co
7							
8		=STDEV(D13:D16)		=(G15^2/B12+G25^2/B22)^0.5			
9							
10	Initial Year	Initial Year	Initial Year	Initial Year			=F25/F15
11							
12							
13	7.090		= Co				
14	7.741						
15	7.378	7.247	1404	0.410			
16	6.780						
17					-1.518	0.236	0.219
18							
19							
20	Final Year	Final Year	Final Year	Final Year			
21							
22							
23	5.620		= Ci				
24	5.961						
25	5.878	5.729	308	0.233			
26	5.455						
27							
28		=STDEV(D23:D26)					
29							
30							

Figure B.2. Calculating the standard deviation of the means, the difference between the means, and the standard deviation of the difference between the means. Calculations are performed on the natural logarithms of the concentration data. Also calculated is an attenuation factor between the mean of samples in the initial year and mean of samples in the final year of the review cycle.

was constructed from a blank spreadsheet, remember to label the column headings as appropriate.

Figure B.1 presents the first data entry cells in the spreadsheet. The sampling dates in the first year go in cells **B13** through **B16** and the dates for the final year go into cells **B23** through **B26**. Enter the number of sampling dates in the first year in **cell B12** and the number of sampling

dates in the final year in **cell B22**. These numbers will be used later to calculate the degrees of freedom in the *t*-statistic. Concentration data for the first year go in cells **C13** through **C16**, and for the final year in cells **C23** through **C26**. Notice that concentrations are in units of $\mu\text{g/L}$. Once this data is entered, the spreadsheet uses the formulas already embedded in the spreadsheet to make a number of calculations.

Enter α , the Acceptable Probability of Error

	J	K	L	M	N	O
1						
2						
3	Attenuation	Probability	Degrees	Critical Value	Difference	Attenuation
4	Factor	of Error	Freedom	Student's	of Means	Factor (Ci/Co)
5	Ci/Co	α one-tailed	in	<i>t</i>	required to be	required to be
6			Student's	(2 α , d.f.)	significant at	significant at
7			<i>t</i>		various levels	various levels
8					of α one-tailed	of α one-tailed
9						
10			4.755			
11					=TINV(2*K12,L\$10)	
12		0.4		0.271	-0.064	0.938
13		0.3		0.569	-0.134	0.875
14		0.2		0.941	-0.222	0.801
15		0.15		1.190	-0.280	0.756
16	0.219	0.1		1.533	-0.361	0.697
17		0.05		2.132	-0.502	0.605
18		0.025		2.776	-0.654	0.520
19		0.010		3.747	-0.883	0.414
20		0.005		4.604	-1.085	0.338
21		0.0025		5.598	-1.319	0.267
22						
23						

=F25/F15

=-I\$17*M12

=EXP(N12)

=((G15^2/B12+G25^2/B22)^2)/((G15^2/B12)^2/(B12-1)+(G25^2/B22)^2/(B22-1))

Figure B.3. Calculating the difference between the means necessary to be statistically significant at a predetermined probability of error. Calculations are performed on the natural logarithms of the concentration data. The differences in means are then back transformed into the attenuations factors that are significant at a predetermined probability of error.

The spreadsheet calculates the natural logarithm of the concentrations in column C, and presents the arithmetic mean of the natural logarithms of the concentrations in the first year and final year of the review cycle in cells **E15** and **E25** respectively (see Figure B.1).

In column F, the spreadsheet calculates the geometric mean of the samples in the initial year (C_o) and final year (C_i) of the sampling interval (Figure B.2). Notice that in the review cycle, there was roughly a four-fold reduction in the concentrations of TCE between the original concentration at the beginning of the interval (C_o) and the concentration at the end of the interval (C_i). The spreadsheet calculates an attenuation factor (C_i/C_o) in cell **J17**.

The next step is to determine whether that extent of attenuation is statistically significant. The spreadsheet will perform calculations to determine whether the difference between the means of the transformed data is significant.

The variance of the difference is simply the sum of the variances of the individual means. The variance is estimated by the square of the standard deviation. In cells **G15** and **G25** the spreadsheet calculates the standard deviation of the transformed data in the first year and in the final year of the review cycle. The difference between the means is calculated in cell **H17** (Figure B.2). Note that the first mean is subtracted from the final mean. If concentrations attenuate, the difference will be a negative number.

In cell **I17**, the spreadsheet calculates the standard deviation of the difference between the means, using equation B.3. In cell **L10** (Figure B.3), the spreadsheet estimates the number of degrees of freedom in t using Equation B.4. Excel 2003 can calculate the critical value of the t -statistic for any probability of error and for any number of degrees of freedom. In Figure B.3, Cells **K12** through **K21** contain example values for the probability of error (α). In Figure B.3, these range from 0.4 to 0.0025. The user can enter any desired

value for α in the spreadsheet. The corresponding level of confidence is $(1 - \alpha)$. If α is 0.05, then the level of confidence is 95%.

In column M, the spreadsheet calculates the critical value of the t -statistic. The function **TINV** calculates the inverse of the t -statistic, which is another way to say the critical value of t .

The formula in Excel for **TINV** has the format **TINV**(*first parameter*, *second parameter*). The *first parameter* is the probability of error and the *second parameter* is the degrees of freedom in the comparison. Excel assumes that the user is making a two-tailed comparison. In other words, half the distribution of error is in values greater than the mean, and half is in values less than the mean. A rate of attenuation that is faster than expected is an acceptable outcome. There is no reason to use statistics to protect against a rate that is faster than expected. We are only interested in rates that are slower than expected. It is more appropriate to assign all of the uncertainty to the slower rate. To use Excel to calculate a one-tailed value of the critical value of the t -statistic, the *first parameter* in the formula for **TINV** must be twice the probability of error that would be expected for a two-tailed test. This is why the formula multiplies the value of α by 2.

Equation B.2 can be rearranged as follows:
 $(\bar{X}_1 - \bar{X}_2) > s_{(\bar{X}_1 - \bar{X}_2)} \times t_{\alpha, d.f.}$ is equivalent to

$(\bar{X}_2 - \bar{X}_1) < -s_{(\bar{X}_1 - \bar{X}_2)} \times t_{\alpha, d.f.}$. Column N calcu-

lates a value for $-s_{(\bar{X}_1 - \bar{X}_2)} \times t_{\alpha, d.f.}$. If the value

of $\bar{X}_2 - \bar{X}_1$ is less (more negative) than the value of $-s_{(\bar{X}_1 - \bar{X}_2)} \times t_{\alpha, d.f.}$ in column N, then \bar{X}_2 and \bar{X}_1

are significantly different at the corresponding level of α in column K.

The actual difference of means $\bar{X}_2 - \bar{X}_1$ was -1.518 (cell **H17** in **Figure B.2**). This difference is significant with a probability of error

less than 0.0025 (compare cell N22 to cells H17 in Figures B.2 and B.3).

The difference of means is the natural logarithm of the attenuation factor (C_i/C_o). The formula in Column O takes the antilogarithm of the differences to recover the attenuation factors that are significant for various levels of α . The attenuation factors in column O are offered to illustrate the capacity of the monitoring data to resolve the extent of natural attenuation over the review cycle. The measured attenuation factor was 0.219 (cell J17 in Figure B.2). As an example, at a preselected probability of error of 0.05, the variation in the samples makes it possible to recognize any attenuation factor that is less than 0.605 over the five year interval in the review cycle (cell O18 in Figure B.3).

Figure B.4 shows data input cells for the final concentration based goal, the time interval involved in the review cycle, and the total time interval from the initial year of review cycle to the date when the clean up goal must be reached. The spreadsheet uses Equation

8 to calculate an interim concentration in the final year of the review cycle that corresponds to each sample in the first year of the review cycle. The interim goals in cells S13 through S16 correspond to the concentrations in cells C13 through C16. The equation in column T calculates the natural logarithm of the interim goals.

As depicted in Figure B.5, the equation in cell U15 calculates an arithmetic mean of the natural logarithms, and the equation in cell V15 recovers the geometric mean of the interim goals. The geometric mean of the interim goals was 241 $\mu\text{g/L}$, compared to a geometric mean of samples in the final year of 308 $\mu\text{g/L}$ (see cell F25 in Figure B.2). The sample mean is larger than the goal. There is a possibility that natural attenuation is not adequate to meet the goal by the time specified.

We will use the same approach with the t -statistic for the difference of means to compare the mean of the interim goals to the mean of the samples in the final year. The equation in

	P	Q	R	S	T
1					
2	Setting Interim Goals (C_{ig}) for Final Year of Review Cycle				
3					
4	Final Goal	Time Interval*	Time Interval	Interim	LN C_{ig} required
5	or MCI	between years	from initial	Goal (C_{ig})	to be adequate
6	($\mu\text{g/L}$)	in review cycle	year to goal	required	to meet goal
7		(years)	(years)	to be on track	
8				to meet	
9				Final Goal	
10				($\mu\text{g/L}$)	=LN(S13)
11					
12					
13	5	5	16	216	5.377
14				339	5.825
15		* length of review cycle		264	5.575
16				175	5.164
17					
18					
19	=C13*(P\$13/C13)^(Q\$13/R\$13)				

Figure B.4. Setting interim clean-up goals for the final year of the review cycle. A separate interim goal is calculated for each sample in the initial year of the review cycle, assuming first order degradation to attain the final clean up goal by the specified date.

	U	V	W	X	Y
1					
2		Comparison of Samples in the Final Year of			
3		to Interim Goals for Final Year of Review			
4	Interim Goals	Interim Goals	Interim Goals	Difference	Stan Dev
5	Mean	Geometric Mean	Stan Dev	of Means	Difference
6	LN Cig	Conc. (µg/L)	LN Cig	(LN Conc.)	of Means
7					
8					
9					
10		=STDEV(T13:T16)			
11					
12	=AVERAGE(T13:T16)				
13					
14					
15	5.485	241	0.282	-0.243	0.183
16					
17	=EXP(U15)		=U15-E25		
18					
19					
20					
21	=(G25^2/B22+W15^2/B12)^0.5				
22					
23					
24					

Figure B.5. Calculating the mean of the interim goals, the standard deviation of the interim goals, the difference between the mean of samples and the mean of the interim goals, and the standard deviation of the difference between the mean of samples in the final year and the mean of the interim goals. Calculations are performed on the natural logarithms of the concentration data or the goals.

cell **W15** calculates the standard deviation of the natural logarithm of the interim goals, the equation in cell **X15** calculates the difference between the mean of the goals and the mean of the samples in the final year of the review cycle, and the equation in cell **Y15** calculates the standard deviation of the difference in the means.

As depicted in Figure B.6, the equation in column AA calculated the degrees of freedom in the comparison of the interim goals to the samples in the final year, and the equation in column AB calculates the critical value of Student's *t*. Note that Excel does not calculate the critical value for a fractional value of the degrees of freedom. Instead it calculates the value for the next lowest integer. The equation in column AC calculates the minimum difference required for the sample mean to

be statistically different from the mean of the interim goals. As the probability of error decreases, the difference of the means become more negative. The actual difference was -0.243 (cell **X15** in Figure B.5). The minimum difference necessary to be statistically significant at a probability of error of 0.15 was only -0.235 (cell **AC16** in Figure B.6). The mean of the interim goals was less than the mean of samples in the final year with a probability of error of 0.15 (cell **Z16** in Figure B.6).

Notice that the minimum difference necessary to be statistically significant at a probability of error of 0.10 was -0.270 (cell **AC17**), which is more negative than the actual difference of the means. There was no evidence that the mean of the interim goals was less than the mean of samples in the final year with a probability of error of 0.10.

$$=((G25^2/B22+W15^2/B12)^2)/((G25^2/B22)^2/(B22-1)+(W15^2/B12)^2/(B12-1))$$

$$=IF(X\$15-AC13<0, "No", "No evidence not adequate")$$

	Z	AA	AB	AC	AD
1					
2					
3					
4	Probability	Degrees	Critical Value	Difference	
5	of Error	Freedom	Student's	of Means	
6	α one-tailed	in	t	required for	
7		Student's	(2α , d.f.)	C_i to be	Attenuation
8		t		statistically	Adequate
9				different	to Attain Goal?
10	=K12			from C_{ig}	
11		5.795			
12					
13	0.4		0.267	-0.049	No
14	0.3		0.559	-0.102	No
15	0.2		0.920	-0.168	No
16	0.15		1.288	-0.235	No
17	0.1		1.476	-0.270	No evidence not adequate
18	0.05		2.015	-0.368	No evidence not adequate
19	0.025		2.571	-0.470	No evidence not adequate
20	0.01		3.365	-0.615	No evidence not adequate
21	0.005		4.032	-0.737	No evidence not adequate
22	0.0025		4.773	-0.872	No evidence not adequate
23					
24			=-AB13*Y\$15		

$$=TINV(2*\$Z13,\$AA\$11)$$

Figure B.6. Evaluation whether attenuation is adequate to attain goals. C_i is the mean of the samples in the final year and C_{ig} is the mean of the interim goals. If the difference of means required for C_i to be statistically different from C_{ig} is less than the difference between C_i and C_{ig} , then the extent of attenuation in concentrations over the review cycle is not adequate to attain the clean up goal in the specified time.

Based on this comparison, we can say with 85% confidence that the trend in natural attenuation over the review cycle from 2001 to 2006 was not adequate to attain the clean up goal. However, we can not make the same statement at 90% confidence.

B.1.4. Modifying the spreadsheet to accommodate samples

The spreadsheet is set up with four samples in the first year and four samples in the final year as a default. If fewer samples or more samples are available, the spreadsheet must be modified. If there are fewer than four samples in the initial year and final year, the pre-existing example values remaining in cells in Column B should

be deleted. If there are more than four samples, the extra samples can be inserted in the blank cells below the existing example data. Insert the correct number of samples in cells **B12** and **B22**. Modify the formulas in cells **E15**, **E25**, **G15**, **G25**, **T21**, **U15** and **W15** to reflect to correct range of the data. To modify a formula, click on the cell containing the formula, and then edit the formula in the *fx* box above the cells of the spreadsheet. You will only edit the second parameter, to reflect the correct range of row numbers in the data set.

As an example, if you only had three samples in the first year, then data would be entered in cells **B13** through **B15** and cells **C13** through **C15**. The equation in cell **E15** would be modified to read **=AVERAGE (D13:D15)**.

Three samples are probably the minimum number of samples in the mean for a comparison of means using the *t*-test. With only two samples, the ability to resolve differences is very low. With more samples, the ability to resolve differences increases. At many sites, one sample each quarter is a reasonable schedule to evaluate natural attenuation, particularly if one sample is collected each quarter in the initial year and final year of the review cycle. However, there are certain hydrogeologic settings such as bare karst terrain with conduit flow where groundwater movement and contaminant concentrations can be extremely variable over short time intervals. For these types of very dynamic hydrogeologic settings, larger sample sizes are typically needed in order to have reasonable assurance that the true population is captured by the sample data. Selection of quarterly sampling for groundwater monitoring is typically done, but there are clearly hydrogeologic settings where other sampling frequencies are more technically justifiable.

B.2. Independence of samples and seasonal effects

A further assumption of the *t*-statistic is that the samples are independent of each other. If replicate samples were taken and analyzed on a particular day, they only count as one sample. You

can average the samples, or randomly choose one sample to enter into the spreadsheet. Wells are usually purged before they are sampled, and the sample represents a composite of the volume of water surrounding the well screen that was produced from the well during purging and sampling. A sample taken on the following day may not be an independent sample, because it could have some component of the water in the previous day's sample. For a sample to be independent, you must allow the ambient flow of the ground water to move all the water that might have been produced in an earlier sample away from the well screen.

One approach is to estimate the seepage velocity of the ground water in the aquifer, and compare the distance ground water travels along a flow path over a time interval to the distance around a well that contributes ground water during a purging and sampling event.

The distance around a well that contributes ground water during a purging and sampling event can be estimated from information on the construction of the well, the volume of water produced during purging and sampling, and an estimate of effective porosity of the filter pack and the aquifer material. The formula for the volume of cylinder can be used to estimate the volume of water contained within a particular radius of capture of the monitoring well (Equation B.5). The equation calculates the volume of water (*V*) that would be contained in aquifer material within a particular radius (r_a). Equation B.5 corrects that volume for the additional water that would be contained in the radius of the filter pack (r_p), and for the additional water that would be contained in the radius of the well itself (r_w). In Equation B.5, (*d*) is the length of the filter pack (not the screen) or the length of the screen if there is no filter pack, and θ_a is the porosity of the aquifer material and θ_p is the porosity of the filter pack if any. If there is no filter pack, assume $\theta_p = \theta_a$.

$$V = \pi r_a^2 d \theta_a + \pi r_p^2 d (\theta_p - \theta_a) + \pi r_w^2 d (1 - \theta_p)$$

B.5

Solving equation B.5 for (r_a) produces Equation B.6.

$$r_a = \left[\left((V/\pi d) - r_p^2 (\theta_p - \theta_a) - r_w^2 (1 - \theta_p) \right) / \theta_a \right]^{1/2} \quad \text{B.6}$$

The diameter of capture of water in the aquifer by the well (D_a) is described by Equation B.7.

$$D_a = 2r_a \quad \text{B.7.}$$

Divide the diameter of capture of water in the aquifer (D_a) by the seepage velocity of ground water in the aquifer to estimate the time that must elapse to be able to collect an independent sample. It would be wise to add a safety factor of a few-fold.

Some sites show strong seasonal effects on concentrations. This may be due to natural heterogeneity or even changes in the elevation of the water table. If you are missing a sample, seasonal effects could add an additional source of uncertainty that is not accounted for in the t -test for the difference of means. If possible, balance the samples across the seasons. If a sample is missing, one alternative to be considered is to shift the entire analysis forward or backward one season to provide a complete sample set for one year of monitoring, even though the “one year of monitoring” is not contained within one calendar year.

B.3. Selecting the appropriate value of α

Faul et al. (2009) provide G*Power 3.1.2, a convenient application that can be used to determine the associated values of α (probability of a Type I error) and β (probability of a Type II error) in a t -test for the difference of means. A screen shot of the input and results screen of the application is provided in Figure B.7. Download and open the application. Open the drop down menu **Test Family** and **select t tests**, open **Statistical Test** and select **Means: Difference between two independent means (two groups)**, open **Type of power analysis** and select **Compromise: Compute implied α and power**

– **given β/α ratio, sample size, and effect size.**

Open the **Determine => Effect size d** submenu and populate the menu. In **mean group 1** enter the arithmetic mean of the natural logarithm of samples in the last year of the review cycle (cell **E25** in the spreadsheet EvalMNA) and in **SD σ group 1** enter the standard deviation of the natural logarithm of the samples (cell **G25** in EvalMNA). Similarly, in **mean group 2** enter the arithmetic mean of the natural logarithm of the interim goals (cell **U15** in EvalMNA) and in **SD σ group 2** enter the standard deviation of the natural logarithm of the interim goals (cell **W15**). Select **Calculate and transfer to main window**. If for some reason, the appropriate ratio of β/α is something other than 1, insert the appropriate value for **β/α ratio**. Then enter the appropriate values for **Sample size group 1** (number of samples in the final year) and **Sample size group 2** (number of samples in the initial year which is equal to the number of interim goals). Select **Calculate**. The **err prob** and **β err prob** are presented in the appropriate windows. For the example data, $\alpha = \beta = 0.256$. This value can be inserted into any cell of column K row 13 to 22 of EvalMNA to evaluate whether the means are different for this precise value of α .

The null hypothesis of the t -test for the difference in means is: The means are not different. The interpretation of the null hypothesis is that the attenuation in concentrations over the time interval in the review cycle is adequate to attain the clean up goal within the time period expected. At this value of α , the t -test for the difference of means rejects the null hypothesis (Figure B.6), indicating that the extent of attenuation is not adequate to meet the long term goals.

Notice that the degrees of freedom in the G*Power 3.1.2 is 6, (Figure B.7) while the degrees of freedom in the comparison in EvalMNA is 5.795, which Excel will truncate to 5 in the calculation of the critical value of Student's t . As a result, for a particular value of α , the critical value of Student's t in EvalMNA

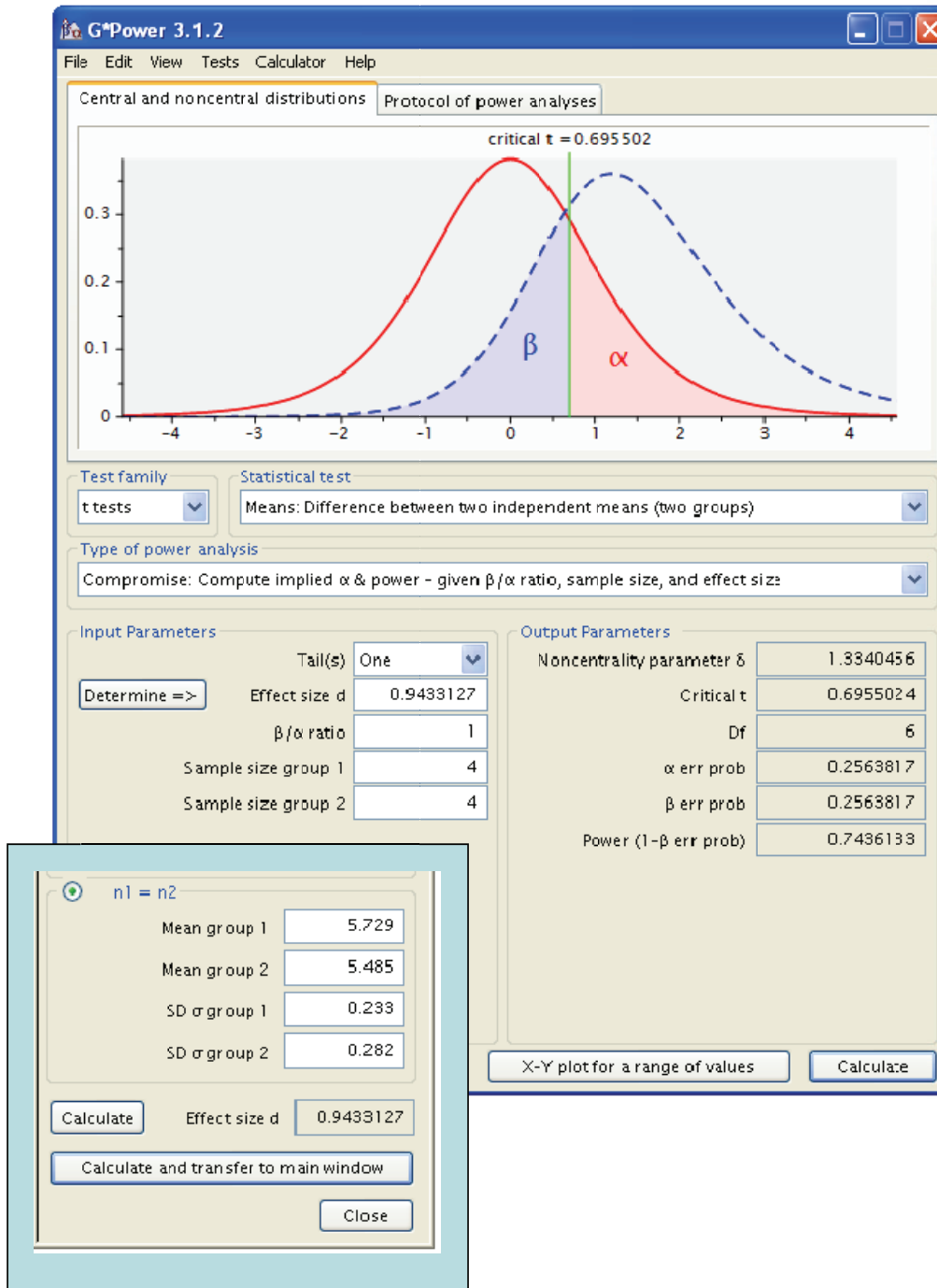


Figure B.7. Data input screen and subscreen for G*Power 3.1.2.

will be a larger number, making the difference in means required for C_i to be statistically different from C_{ig} a larger number. For $\alpha = 0.256$, the corresponding critical value of t in G*Power 3.1.2 is 0.696 compared to the critical value of t in EvalMNA of 0.706. The critical value of t in EvalMNA is conservative in that more attenuation will be required before the statistical analysis will determine that attenuation is not adequate to attain the goal.

Reference:

Faul, F., E. Erdfelder, A. Buchner and A.-G. Lang. 2009. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses, *Behavior Research Methods*, 41 (4), 1149-1160, doi:10.3758/BRM.41.4.1149. Both the journal article and the computer application are available at <http://www.psych.uni-duesseldorf.de/aap/projects/gpower/>.

ISSUE



PRESORTED STANDARD
POSTAGE & FEES PAID
EPA
PERMIT NO. G-35

Office of Research and Development (8101R)
Washington, DC 20460

Official Business
Penalty for Private Use
\$300