Wisconsin
Department of
Natural Resources

Bureau for
Remediation and
Redevelopment

PUBL RR-523-03
October 7, 2003

Guidance on the Use of Leaching Tests for Unsaturated Contaminated Soils to Determine Groundwater Contamination Potential

Go to Addendum with Examples

Purpose

This guidance pertains to leaching tests conducted on contaminated soils, to determine the potential for groundwater contamination exceeding the standards in ch. NR 140 - Groundwater Quality Standards, Wis. Adm. Code. This guidance is not intended to be applicable to leaching tests conducted to determine waste characteristics or the stability of wastes that have been physically or chemically stabilized. This guidance addresses leaching tests used to establish site-specific residual contaminant levels according to NR 720.19(4)(b)2, Wis. Adm Code.

Disclaimer:

This document is intended solely as guidance, and does not contain any mandatory requirements except where requirements found in statute or administrative rule are referenced. This guidance does not establish or affect legal rights or obligations, and is not finally determinative of any of the issues addressed. This guidance does not create any rights enforceable by any party in litigation with the State of Wisconsin or the Department of Natural Resources. Any regulatory decisions made by the Department of Natural Resources in any matter addressed by this guidance will be made by applying the governing statutes and administrative rules to the relevant facts.

This guidance is based on requirements found in chs. NR 140, 720, 722, 724, and 726, Wis. Adm. Code; the Hazardous Substance Spill Law, s. 292.11, Wis. Stats., the Environmental Repair Statute, s. 292.31, Wis. Stats., and the Groundwater Law, ss. 160.23 and 160.25, Wis. Stats.

OTHER RELEVANT GUIDANCES

This guidance may be more complete when used in conjunction with:

Understanding Wisconsin Standards for Cleanup of Contaminated Soil, An Overview of Wisconsin Administrative Code Chapter 720, Publication RR-520-97
Interim Guidance on Soil Cleanup Levels for Polycyclic Aromatic Hydrocarbons (PAHs), Publication RR-519-97
Interim Guidance for Selection of Natural Attenuation for Groundwater Restoration and Case Closure Under Section NR 726.05(2)(b), Publication RR-530-97
Interim Guidance on Soil Performance Standards, Publication RR-528-97

This list will be expanded as additional guidances are developed.

These guidances may be obtained by:

A. Sending a request to: Public Information Requests, Bureau of Remediation and Redevelopment, Department of Natural Resources, P.O. Box 7921, Madison, WI 53707.

B. Calling the publication coordinator at (608)261-4391.

This guidance will be updated as needed. Comments and concerns may be sent to "Guidance Revisions", Dale Ziege - RR/3, WDNR, P.O. Box 7921, Madison, WI 53707.

THE USE OF LEACHING TESTS FOR UNSATURATED CONTAMINATED SOILS TO DETERMINE GROUNDWATER CONTAMINATION POTENTIAL

INTRODUCTION: Leaching tests are one of the methods identified in s. NR 720.19(4)(b) as appropriate for use in developing site-specific soil standards for the protection of groundwater. The use of leaching tests to simulate the potential for contaminants to leach from soils to the groundwater is very promising. This guidance tries to identify the key issues associated with the use of leaching tests for developing site-specific soil standards for groundwater protection. Also, an attempt is made to identify significant trouble spots with the use of leaching tests.

TYPES OF LEACHING TESTS: There are two main categories of leaching tests:

1. Batch tests -- Where contaminated media is sampled and a known mass placed in a vessel with a known volume of leachant (extracting fluid). The vessel is then agitated by tumbling or shaking for a set time, to complete the extraction.

The major limitations of this type of test is that it does not represent actual field conditions and may over or under estimate the mass of contaminant that would leach from a specific mass of contaminated soil.

2. Column tests -- Where contaminated media is sampled by obtaining an intact core of soil that remains in a cylinder, or by repacking a column of soil into a cylinder. A known volume of leachant is passed through the column.

The major limitation with this type of test is that, while it attempts to simulate field conditions, it is virtually impossible to determine whether preferential pathways have formed in the column. Therefore, the amount of contaminant leached may grossly underestimate the actual contaminant mass that may leach from the soil.

The advantage of using leaching tests is that the test can be conducted on soils from the site. Theoretically, the leachate from the leaching test should represent the actual amount of contaminant that will leach from the soil minus the dilution from the extracting fluid.

Both the batch test and the column test have inadequacies for determining the actual potential for a contaminant to migrate to the groundwater. In general, the common limitation of leaching tests is that only the amount of contaminant that will leach from a set mass of contaminated soil under specific conditions is determined. This information should be used in a way that is appropriate for the purpose, identified in the objectives for conducting the leaching test. Before applying the data from a leaching test, the mechanisms controlling contaminant leaching from the soil should be identified and understood. The release of contaminants from soils is extremely complex. An evaluation of factors that may influence the results of the leaching test should be performed. Some tests use aggressive leachant (pH of 2), as in the toxic characteristic leaching procedure

(TCLP), to simulate the conditions found in the core of a landfill. The TCLP test could underestimate the amount of contaminant leached by immobilizing constituents that would be mobile under normal soil conditions. The TCLP could overestimate the amount of contaminant mass by leaching constituents that would not be mobile under normal soil conditions. The TCLP is not recommended for use in determining site specific RCLs. Leaching tests also vary on the duration of the leaching procedure; some tests last days, weeks, even months. A good source of scientific considerations that should be taken into account prior to applying leaching tests outside of this guidance is "Leachability Phenomena. Recommendations and Rational for Analysis of Contaminant Release by the Environmental Engineering Committee" (U.S. EPA 1991).

USING LEACHING TESTS TO ESTABLISH SITE-SPECIFIC RCLs UNDER NR 720

When Leaching Tests Can Be Used: As provided in s. NR 720.05(1), Wis Adm. Code, leaching tests can only be used to establish site-specific RCLs after an investigation has been conducted and the site has been fully characterized and the degree and extent of contamination have been determined. Leaching tests will generally be worthwhile only after a remedial action has been completed, or when the levels of contaminants determined in the site investigation are not expected to cause a groundwater threat. A leaching test would be used to determine the potential for the remaining contaminants to desorb from the soil and contaminate ground water to a level that may exceed the groundwater standards. Leaching tests are most applicable to determining residual contaminant levels protective of groundwater for soil contaminated with heavy metals. It is difficult to use a fate and transport model to simulate the movement of heavy metals through unsaturated soils, due to the complexity of both the chemical and physical properties of soils and the contaminants. However, leaching tests can also be used for inorganic and organic contaminants, both volatile and non-volatile.

Sample Collection and Handling:

1. Sampling Plans:

Sampling of soils for use in a leaching test should follow the same protocol used for determining types and numbers of samples for any soil sampling plan. Both of the following are good references for help in developing proper sampling plans: "Methods for Evaluating the Attainment of Cleanup Standards - Volume 1: Soil and Solid Media" and "Statistical Methods for Environmental Pollution Monitoring." The number and types of samples should be determined by the expected use of the data. The number of samples will also depend on the application of statistics to interpret the degree of confidence in the data. It is suggested that the EPA's Data Quality Objective process be used in determining the soil sampling plan. (See Table 1 for an example of this process.)

In general, samples should be collected from areas representative of the contamination at the site and should adequately address soil heterogeneity at the site. Variations in soil are difficult to standardize as described by L.P. Wilding (1985):

"Spatial variability of soil is not an academic question. It is a real landscape attribute; our unwillingness or inability to identify it in no way decreases its magnitude or existence. ... As scientists we must document the magnitude and form of soil variability; accommodate its existence in models of soils; and transmit accurately the expected pattern and implication of spatial changes to users of soil resources. Soils are not material specific; many soil properties are not single valued, many are transient, and many are not randomly distributed but rather systematically time and spatially dependent. The dilemma is that soils are not isotropic media but rather they are strongly anisotropic laterally and vertically."

2. Sample Collection:

Sample collection protocols are dependent on the type of contaminant to be sampled, the purpose for which the samples are collected and the type of media to be sampled. This guidance pertains to the collection of soil samples to be used in leaching tests. The following sections discuss considerations for sample collection for that purpose:

a. Metals, non-volatile organics and inorganic containments:

Samples should be collected and handled similar to samples being collected for routine analysis. Care should be taken to assess the conditions of the soil sample and the possibility for changes that may occur due to chemical or biological processes. Leaching tests with acidic leachants should not be used for certain metallic ions with multiple valance states that behave as an anionic species. Leaching procedures that add acidic solutions may convert the anion to a cation and this would cause an incorrect measurement of the actual leaching potential of the contaminant. For example, hexavalent chromium will react with the acidic extracting fluid of some leaching tests to form trivalent chromium, a less mobile form of chromium. Therefore, that leaching test may underestimate the amount of chromium that would leach from soils contaminated with chromium.

b. Volatile contaminants:

When sampling volatile organic compounds (VOCs), the contaminants may volatilize before the test is conducted. Therefore, care must be taken to prevent volatilization by using a device which insures sample integrity as required in s. NR 700.13 Wis. Adm. Code, **other than methanol preservation (which would increase the leaching of VOCs)**. Neither the sample that is to be used in the leaching test nor the totals sample for comparison should be preserved in methanol. The use of brass tubes with capped ends, or the Encore® sampler or similar sampling device, is needed to accurately represent the contaminants present at the site. To limit volatile losses, the time between sample collection and sample testing should be minimized and should be considered during the sampling plan design. The holding times for the brass tube is two hours and for the encore sampler is forty-eight hours. This makes good coordination with the laboratory critical to the success of the method. Things to keep in mind when using a leaching

test for volatile organic compounds are: laboratory holding times (holding times are dependent on sample collection methods), the number of Zero Headspace Vessels the laboratory has available for the procedure (usually not more than 10 vessels), the extracting time is 18 hours (time of day the samples arrive determines when it can be run), and the total volatile sample should be run at the same time the extraction begins, to allow comparisons between leached and total VOCs on a sample. If a sample is analyzed to determine if a leaching procedure will be conducted for volatile compounds, that same sample can not be used for the leaching procedure due to the variable holding times.

Suggested Test: The test procedure that is recommended for use in determining soil contaminant leaching potential is EPA Method 1312, the synthetic precipitation leaching procedure (SPLP). The SPLP can be used for both organic and inorganic contaminants. Soils contaminated with VOCs can be analyzed using this method, by using the zero headspace vessel option. This test is essentially a single-point batch desorption experiment with standard solutions, preparation of the materials, soil:solution ratio, and equilibration time being established. The preparation of the extraction solution is specified and is to be pH 4.2 +/- 0.05 for materials collected east of the Mississippi River.

As discussed above, the SPLP test is a single-point batch desorption procedure. This test was designed to determine the mobility of both organic and inorganic analytes present in samples of soils, wastes, and wastewaters. The test itself is written so that standard procedures are followed, but there are several factors that may be of concern when applying this test to contaminated soil and aquifer materials.

In the SPLP test, if the particle size is greater than a specific size, the particle size is reduced. Reduction of the particle size would increase the surface area of the sorbent and possibly increase the sorption of the contaminant. This is particularly important for sorbents low in organic carbon content and for solutes that partition due to mechanisms other than hydrophobic partitioning, such as ion exchange. Exposure of the sample to the atmosphere when reducing the particle size by crushing, cutting, or grinding may cause loss of volatile organic components or of residual NAPLs, thereby changing the partition coefficient as well as losing the contaminants of interest. If the sample is dried, the oxidation of the natural organic carbon will change and its solubility increase, which will affect the measured partition coefficient.

A standard equilibration period of 16-20 hours is recommended, which is a common equilibration period for many tests. Many systems, especially metals, do not reach equilibration within 24 hours and desorption will be underestimated. More information about the solute-sorbent interactions can be obtained by also conducting a kinetic experiment where samples are collected after various times, throughout the equilibration period, which may range from 2 to 96 hours or more.

The solution is separated from the sample by filtration rather than centrifugation. Filtering is an additional step where some of the components could be lost due to volatilization and to exposure

to the filter and filtration apparatus. For standard soil materials, centrifugation may be preferable, although care must be taken for materials contaminated with a residual NAPL phase which is less dense than water.

The use of site-specific materials in a standard test, such as the SPLP, will not necessarily guarantee that more accurate values are obtained for the desorption of contaminants. Therefore, without significant scientific documentation and validation, site specific sorption values that can be determined using leaching tests should not be used in other contaminant and fate models.

It should be noted that the SPLP test inherently has a 20:1 dilution factor. This dilution factor is the only dilution that should be used, unless a much more extensive scientific analysis of the controlling mechanisms of leachability phenomena is conducted at the specific contaminated site.

In addition to the limitations previously discussed, on losses from volatilization, degradation, and laboratory scheduling, there is the possibility of analytical difficulties. These problems include, but are not limited to, differences between the number of contaminants in the total sample and the leached sample, method detection limit and quantification limit issues. Many of these issues can be standardized by the development of a good standard operating procedure (SOP) for the analytical procedures used in obtaining data from leaching tests. It is recommended that consultants review their laboratory's SOP for the SPLP leaching tests.

Use of Leaching Test Data: It is recommended that leaching test data be used in one of the following ways:

A. The leachate concentration measured by the SPLP test should be compared directly to the preventive action level (PAL). The first step in determining this is to collect samples of the contaminated soils, using random sampling. The total concentration of contaminant in the soils is to be compared to the leached concentration of contaminant for the purpose of determining the distribution of the samples. Using a statistical analysis, determine if the leaching test data shows that the upper 95% confidence limit on the mean of the samples taken at the site will not exceed the PAL. The first step in this analysis is to determine if the data is normally or log-normally distributed. The data can be graphed to determine its distribution or one of many assessments of normality can be performed. The minimum number of samples to determine this is three samples, if the three samples are distributed in a normal or log-normal fashion. In most cases, it will take more than three samples to accurately determine the distribution. If the soils at the site are heterogeneous, a larger number of samples should be collected to account for the potential for greater variability. The samples should be representative of the residual contamination at the site. If the totals sample shows contamination but the leached sample shows no contamination, the value that should be used in the statistical analysis should be one half the limit of quantitation. **All samples should be used in the statistical analysis unless there are quality assurance quality control (QAQC) reasons not to use the data. An increase in variability of the data is not an acceptable reason to ignore a data point.** After the data has been

confirmed to be properly distributed the 95% confidence interval should be determined. Remember to log transform the data, prior to doing the statistical analysis. The use of confidence limits will help compensate for the number of samples analyzed and the variability between samples; the fewer the number of samples the wider the confidence limits on the mean. The department recommends using the 95% confidence limit on the mean, which means that there is a 1 out of 20 chance that the true mean is a higher value and the PAL will be exceeded.

If the upper 95% confidence limit of the mean yields a concentration higher that the PAL, the leaching test should not be used to determine site specific RCLs to determine compliance with ch. NR 720 soil standards because the residual soil contamination may cause a groundwater standard exceedance. If the upper 95% confidence interval yields a concentration lower then the PAL, for all soil units, the residual contamination at the site is not expected to cause a groundwater standard exceedance.

B. An analysis of leaching test data can also be done by determining the mean ratio of the total soil concentration to the leached concentration. Theoretically, the ratio of total contaminant in the soil to the leached concentration should be the same for all samples of similar soil types. If the site is homogeneous, simple random sampling should be used. To sample a site with heterogeneous soils, stratified random sampling should be used. Stratified random sampling is appropriate for heterogeneous soil, where different soil types can be broken down into internally homogeneous units. After the mean of the ratios of the total soil concentration to the leached concentration is determined, the 95% confidence limit should be determined. The value at the lower 95% confidence limit should be determined. Using the value determined above and following equation, the ratio can be used to determine the residual contaminant level:

$$SSRCL = \frac{C_s}{C_l}(PAL)$$

SSRCL - site-specific residual contaminant level (µg/kg)
C_s - Total concentration in the soil (µg/kg)
C_l - concentration leached from soil (µg/l)
PAL - preventive action limit (µg/l)

This equation basically states that for a given ratio of total concentration to leached concentration by using the preventive action limit, the residual contamination left in the soil would not cause a groundwater exceedance.

On a case by case basis with proper scientific documentation and validation, leaching test results can be used in ways other than the two methods discussed above. In the future, as additional information becomes available and test protocols change the department will update this guidance.

REFERENCES:

U.S. EPA 1991. "Leachability Phenomena. Recommendations and Rational for Analysis of Contaminant Release by the Environmental Engineering Committee." EPA-SAB-EEC-92-003. Science Advisory Board, Washington DC.

U.S. EPA 1992. "Preparation of Soil Sampling Protocols: Sampling Techniques and Strategies." EPA/600/r-92/128

U.S. EPA 1994. "Guidance For The Data Quality Objectives Process." EPA QA/G-4

Wilding, L.P. 1985. "Spacial Variability: It's Documentation, Accommodation, and Implication to Soil Surveys." In Neilsen, D.R. and J. Bouma, eds. Soil Spacial Variability. PUDOC. Wageningen, the Netherlands. pp. 166-193

SUGGESTED STATISTICAL REFERENCES:

Gilbert, Richard O., 1987. Statistical Methods for Environmental Pollution Monitoring. Van Nostrand Reinhold, New York, NY. 320pp.

Gonick, Larry., and Smith, W., 1993, The Cartoon Guide to Statistics. Harper Perennial, New York, NY. 230pp.

U.S. EPA. 1989. Methods for Evaluating the Attainment of Cleanup Standards - Volume 1: Soil and Solid Media, U.S. EPA Monitoring Systems Laboratory. Las Vegas, NV.

TABLE 1 - EPA's DATA QUALITY OBJECTIVES PROCESS

DQO Process	Decisions
STATE THE PROBLEM	
• Develop conceptual site model	Identify and document sources and types of contaminants, contaminated media, migration pathways, potential physical targets or receptors.
• Define exposure scenarios	Verify residential land use, other possible exposure pathways, potential for ecological impacts.
• Specify available resources	Identify sampling and analysis budget, scheduling constraints, available resources.
IDENTIFY THE DECISION	
• Identify decision	Determine whether the result of leaching test for particular contaminants exceed the groundwater standard.
• Identify alternative actions	Eliminate the need for further analysis; or use alternative method, i.e. fate and transport modeling.
IDENTIFY INPUTS TO THE DECISION	
• Identify inputs	Specify spacial dimensions of field investigation and delineate on site map.
• Define basis for action level	Remaining contaminated soil produces leachate that exceeds groundwater standard.
• Identify analytical methods	List feasible analytical method consistent with program requirements.
SPECIFY THE STUDY BOUNDARIES	
• Define geographic areas of field investigation	Specify spacial dimensions of field investigation and delineate on site map.
• Define population of interest	Remaining contaminated soil.
• Define scale of decision making	Exposure area (EA) is 0.5 acre for residential land use.
• Identify practical constraints	Identify potential impediments to sample collection, such as access, soil composition, etc.

TABLE 1, continued

• Subdivide site into EAs.	Subdivide each region into 0.5-acre EAs. Delineate site regions where contaminant patterns are likely to be different, based on prior sampling, or etc.
• Define temporal boundaries of study	Identify all known temporal cyclical variations in contaminants; define sampling schedule.
DEVELOP A DECISION RULE	
• specify parameter of interest	Leachate concentration
• Specify a compliance level	No groundwater contamination above standard.
• Specify "if..., then..." decision rule	If the leachate concentration exceeds the preventive action limit (PAL), then determine soil cleanup standard using a different method.
SPECIFY LIMITS ON DECISION ERRORS	
• Define baseline condition (null hypothesis)	The site is contaminated and will cause groundwater contamination.
• Define both types of decision errors	Type 1 (false positive) Will not contaminate groundwater above standards (cause groundwater contamination) Type 2 (false negative) Do more work on the site (site will not cause groundwater contamination)
• Define gray region	Gray region from 1/2 the groundwater standard to 2 times the standard.
• Assign acceptable probabilities of Type 1 and Type 2 decision errors	Type 1: 0.05 (5%); Type 2: 0.20 (20%).

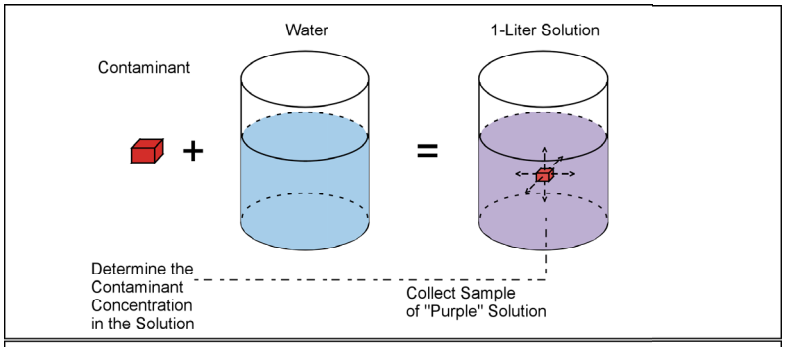
Understanding Synthetic Precipitation Leaching Potential (SPLP), Its Limitations and Data Analysis: *Addendum to PUBL RR-523-03*

Abstract. This addendum to the "SPLP guidance" is specific to *volatile* contaminants. It is intended to help DNR staff, consultants and other interested parties interpret SPLP data when determining statistically-robust residual contaminant levels (RCLs) for volatiles that would be protective of groundwater.

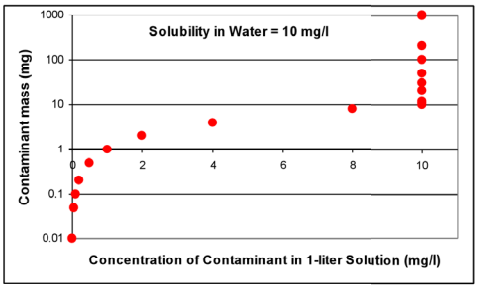
INTRODUCTION

Basic Chemistry. Water is the solvent/extractant used in the SPLP test on volatiles. There is a limit to how much (of the contaminant mass) water can dissolve. For a single contaminant, this limit is the contaminant's solubility in water (S_w).

Illustration. Let's make a 1-liter solution containing a contaminant with a S_w of 10 mg/l.



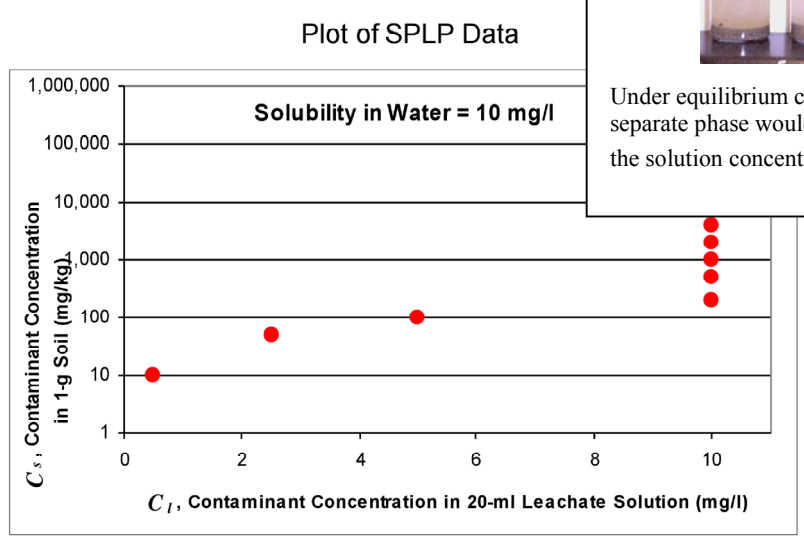
Mass of Contaminant (mg)	Contaminant Concentration in the Solution (mg/l)
0.01	0.01
0.05	0.05
0.1	0.1
0.2	0.2
0.5	0.5
1	1
2	2
4	4
8	8
10	10
11	10
12	10
20	10
30	10
50	10
100	10
200	10
1000	10

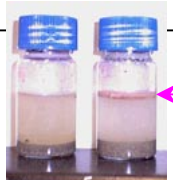


When the solution concentration is at S_w, we expect no further change in solution concentration, even with the addition of more of the contaminant into the mixture.

Relevance to SPLP. The process involved in the SPLP for a *volatile* compound in soil is much like the illustration above, except for the following: 1.) The contaminant mass would be expressed as a concentration in a soil sample; and 2.) The volume of water would typically be less than 1 liter.[1] We can scale the "data" in the illustration, as they would appear in SPLP analyses. First, let's assume zero soil adsorption and that the given contaminant mass is in 1 g of soil (so we divide the contaminant mass by 1 g to express it in mg/g and then multiply the quotient by 1000 g/kg to express it in units of mg/kg). Second, the 20:1 SPLP dilution would require for a 1-g of soil sample only 20 ml of water. With only 20 ml of water, less actual contaminant mass can be dissolved, so we scale the above illustration's solution concentrations by the factor 50 (= 1000 ml / 20 ml), and when the result is more than 10 mg/l (contaminant's S_w), we simply use 10 mg/l. We show the result of our scaled experiment below.

Concentration of Contaminant in 1-g Soil Sample (mg/kg)	Maximum Contaminant Concentration in 20-ml Leachate (mg/l)
10	0.5
50	2.5
100	5
200	10
500	10
1,000	10
2,000	10
4,000	10
8,000	10
10,000	10
11,000	10
12,000	10
20,000	10
30,000	10
50,000	10
100,000	10
200,000	10
1,000,000	10





Under equilibrium conditions, a separate phase would form, keeping the solution concentration at S_w.

It should be obvious from the illustration that low soil concentration data are preferable in establishing any SPLP "trend."

[1] The SPLP method 1312 (http://www.epa.gov/epaoswer/hazwaste/test/pdfs/1312.pdf) describes using a minimum of 100 g of soil sample (and hence 2000 ml or 2 l of water). However, discreet soil samples that are submitted for SPLP test typically weigh less than 100 g. For instance, Encore® samples can weigh about 25 g (http://www.ennovativetech.com/encore/encore.htm), so about 500 ml of water would be used. Laboratories scale the SPLP procedure to accommodate this limitation. The example (soil sample that weighs only 1 g) exaggerates the fact that saturation is reached at a lesser contaminant concentration when a small sample is used in the procedure. The water:soil ratio of 20:1 may not be enough for most volatiles (*e.g.* Roy *et al.*, 1992) – yet another reason *not* to include soils with high contaminant concentrations.

DATA ANALYSIS OBJECTIVE

SPLP Data. The SPLP test was designed to determine whether a soil contaminant is mobile or not - essentially to show whether a contaminant is detected in the solution or not. When done for a number of contaminated soil samples with varying amounts of the contaminant, an extension of the test's utility is in the analysis of the paired data (dry soil contaminant concentration C_s *and* the corresponding SPLP leachate's concentration C_l). Through the analysis of C_s *and* C_l we can empirically estimate a soil contaminant concentration (C_s) below which we do not expect the "leachate" concentration (C_l) to exceed the contaminant's NR 140 preventive action level (*PAL*). In other words, we can determine a residual contaminant level (RCL) protective of groundwater to the *PAL*. In the SPLP guidance (PUBL RR-523-03, p. 8), the RCL "predictor" is through the statistical analysis of the observed ratios C_s / C_l. What should be obvious from the previous illustrations is that S_w dictates where this predictor (C_s / C_l) would fail because there is a limit to C_l (*i.e.* $C_l \leq S_w$), even when apparent C_s keeps increasing. When the leachate concentration $C_l = S_w$, the contaminant concentration in the soil (C_s) becomes irrelevant in establishing any purposeful trend. The empirical predictor C_s / C_l would be biased high (hence erroneously resulting in a high RCL) when SPLP data on soils with very high soil contaminant concentrations are used in the analysis. What we need is a cut-off concentration to avoid this bias when estimating the RCL.

High Concentration Cut-Off for SPLP Data Analysis. A definite cut-off concentration for the soil data would be at a concentration near the soil saturation limit, C_{SAT}, when the contaminant concentration in the soil-water would be at S_w. (The analysis of the SPLP data to determine an RCL should exclude data with soil concentrations near C_{SAT} and greater.) Furthermore, since the "area" of our interest lies near the NR 140 *PAL*, the result of the statistical analysis of SPLP data would be more robust when we analyze data within the vicinity where we expect the leachate concentration to be at the *PAL* level. Briefly, we approximate a range of soil contaminant concentrations (C_s) that can provide a robust *PAL*-based RCL from SPLP data by:

$$0 < C_s < 100 \frac{PAL}{S_w} C_{SAT} \qquad \text{[Important constraint: } 100\,PAL < S_w\text{]} \qquad \textbf{(1)}$$

The cut-off soil contaminant concentration from SPLP test to determine soil-to-groundwater RCLs in equation (**1**) points to the importance of: 1.) Estimating C_{SAT}; and 2.) Using the PAL / S_w ratio in filtering the data before the statistical analysis. The factor 100 in the equation would need to be <u>reduced</u> when the ratio S_w / PAL is less than 100. With $PAL \ll S_w$ for almost all volatiles with NR 140 standards, the data at the low-end soil concentration will be more meaningful in determining the soil-to-groundwater *RCL*.

EXAMPLE CALCULATION TO DETERMINE CUT-OFF SOIL CONTAMINANT CONCENTRATION

Let's estimate the soil saturation concentration C_{SAT} for *TCE*.

Chemical Parameters for TCE: 1,100. 166. 0.422

$$C_{SAT}(mg/kg) = \frac{S_w}{\rho_b}\left(K_{oc} f_{oc} \rho_b + \theta_w + H'\theta_a\right) \qquad \textbf{(a)}$$

Soil Parameters: 1.5 0.20 0.23

Where:

Soil parameters:	Chemical-specific parameters for *TCE*:
ρ_b = dry bulk density = 1.5 g/cm^3	
f_{oc} = fraction of organic carbon	K_{oc} = partitioning coefficient = 1.66e+02 ml/g
θ_w = water-filled porosity = 0.20	H' = dimensionless Henry's law constant = 4.22e-01
θ_a = air-filled porosity = n - θ_w	S_w = 1.10e+03 mg/l
n = total porosity = θ_w + θ_a = 0.434	

Plugging the tabulated values into (a), we get the C_{SAT} for *TCE* as a function of f_{oc}:

$$C_{SAT}(mg/kg) = 182{,}600\ (f_{oc}) + 218 \qquad \textbf{(b)}$$

When f_{oc} = 0.02 (or 2%), C_{SAT} *(mg/kg)* = 3.9e+03 (rounded to 2 significant figures).

The contaminant TCE has an NR 140 *PAL* of 0.5 ug/l (= 0.0005 mg/l). The ratio S_w / PAL = 2.2e+06, so the factor 100 in equation (**1**) is adequate for our purpose.

When the SPLP data set is analyzed to come up with an RCL for TCE that is protective of groundwater, equation (**1**) approximates the useful range of soil concentrations to be:

$$0 < C_S < 0.18 \ mg/kg$$

Naturally, a site with an f_{oc} larger than 0.02 would have a higher cut-off concentration for inclusion in the analysis (whereas a site with f_{oc} < 0.02 would have a lower cut-off).

In the next section, we will use the value of 0.18 mg/kg (= 180 ug/kg) to pre-filter an *example* SPLP soil data on TCE before we perform a statistical analysis to determine the *PAL*-based RCL.

DATA ANALYSIS

Steps to Estimate RCL from SPLP Data and Their Empirical Basis

1.) Determine f_{oc}.
Use the geometric mean of the f_{oc} data when they vary by a factor of 10 or more.

2.) Estimate C_{SAT}.

3.) Determine the range of soil contaminant concentrations using equation (**1**).
Do not include SPLP data above the high-concentration cut-off for subsequent statistical analysis. [Note on non-detect results: Avoid using the data pair (C_s, C_l) when the leachate C_l is ND or below the detection level. When the leachate C_l is quantified, but the dry soil C_s is ND, use 50% of the soil detection limit as the value for C_s.]

4.) Plot total soil concentration C_s (along the y-axis) and leachate concentration C_l (x-axis).

5.) Use the Freundlich isotherm equation:

$$C_s = K\, C_l^{\,n} \qquad (\mathbf{2})$$

Equation (**2**) <u>empirically</u> relates dry soil concentration C_s and leachate concentration C_l [*e.g.* Domenico and Schwartz, 1990, p. 441; Zheng and Bennett, 2002, p. 81] to estimate the parameters K and n.

6.) Once K and n have been determined, then:

$$RCL = K\,(PAL)^n \qquad (\mathbf{3})$$

Important: Both K and n would need to be non-negative numbers.

Alternative Step

7.) A "simplification" to steps 5 and 6 above is to assume that $n = 1$, and K is determined by the lower 95% confidence limit of the mean (LCL) of the C_s / C_l ratios. This step is straightforward after the LCL is determined:

$$RCL = (LCL \text{ of } C_s / C_l)\, PAL \qquad (\mathbf{4})$$

Details for Steps 5 & 6 to Determine Freundlich Parameters K and n. We first describe one strategy where the parameters *K* and *n* are determined by "least-squares" using the *appropriate* SPLP data.

Data

Sample	Total TCE (ug/kg)	SPLP TCE (ug/l)
B-18	175.0	35.0
B-22(1-3)	480.0	110.0
B-22(4-5)	1,100.0	180.0
B-22(6-7)	3,000.0	0.7
B-23	29,000.0	68.0
B-23(6-7)	91.0	2.0
B-26(10-11)	72.0	2.0
B-26(1-3)	85,000.0	3,500.0
B-26(6-7)	750.0	30.0
B-27	640.0	34.0
B-27(6-7)	3,600.0	88.0
B-29	2.0	0.1
B-29(1-3)	220.0	6.2
B-37(1-3)	8.0	1.0
B-37(4-6)	15.0	1.7
B-37(5-7)	50.0	5.0
B-38(1-3)	1,600.0	45.0
B-38(5-7)	20.0	1.9
B-41	94.0	2.5
B-41(1-3)	2,300.0	0.7
B-41(5-6)	73,000.0	0.9
B-45	130.0	60.0
B-45(8-10)	170.0	57.0

Data Sorted

Sample	Total TCE (ug/kg)	SPLP TCE (ug/l)
B-29	2.0	0.1
B-37(1-3)	8.0	1.0
B-37(4-6)	15.0	1.7
B-38(5-7)	20.0	1.9
B-37(5-7)	50.0	5.0
B-26(10-11)	72.0	2.0
B-23(6-7)	91.0	2.0
B-41	94.0	2.5
B-45	130.0	60.0
B-45(8-10)	170.0	57.0
B-18	175.0	35.0
B-29(1-3)	220.0	6.2
B-22(1-3)	480.0	110.0
B-27	640.0	34.0
B-26(6-7)	750.0	30.0
B-22(4-5)	1,100.0	180.0
B-38(1-3)	1,600.0	45.0
B-41(1-3)	2,300.0	0.7
B-22(6-7)	3,000.0	0.7
B-27(6-7)	3,600.0	88.0
B-23	29,000.0	68.0
B-41(5-6)	73,000.0	0.9
B-26(1-3)	85,000.0	3,500.0

Using 0.18 mg/kg (or 180 ug/kg) as the cut-off concentration, we obtain the subset which we would further analyze:

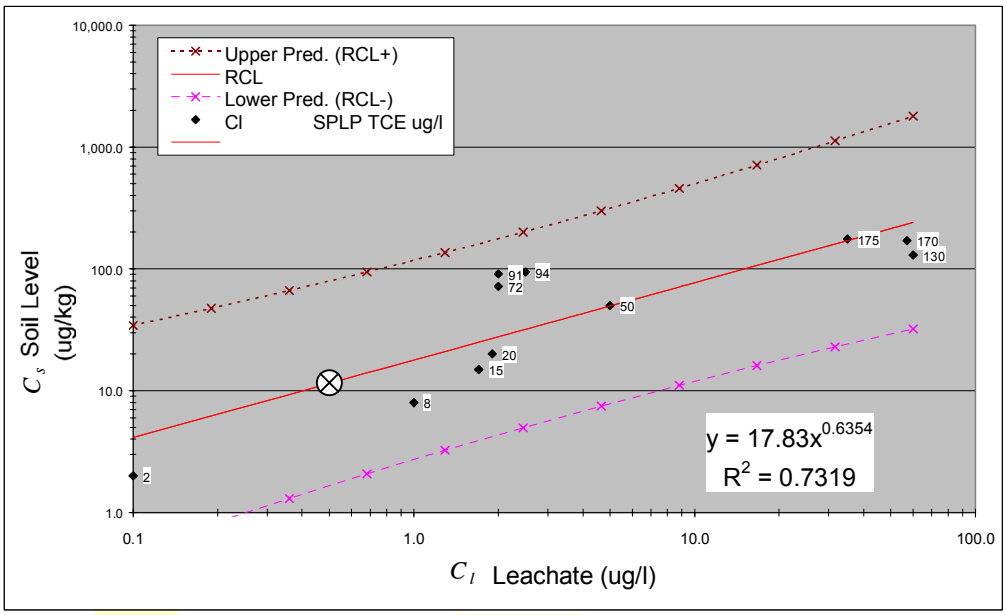
	y (C_s)	*x* (C_l)
Sample	Total TCE (ug/kg)	SPLP TCE (ug/l)
B-29	2.0	0.1
B-37(1-3)	8.0	1.0
B-37(4-6)	15.0	1.7
B-38(5-7)	20.0	1.9
B-37(5-7)	50.0	5.0
B-26(10-11)	72.0	2.0
B-23(6-7)	91.0	2.0
B-41	94.0	2.5
B-45	130.0	60.0
B-45(8-10)	170.0	57.0
B-18	175.0	35.0

Our analysis is summarized in the plot (next page) where we have used Microsoft Excel's *Chart* to *Add Trendline* on the plot of *x*-*y* data. The particular trendline type is the *Power* option that uses the fitting equation $y = c\,x^b$. The Excel regression yields a *c*-value for *K* (= 17.83 for the above data set) and a *b*-value for *n* (= 0.6354). We can then use *K* and *n* from the resulting trendline equation to determine what we are after: $RCL = K\,(PAL)^n$.

With equation (**3**) and given a *PAL* of 0.5 ug/l (our *x* value), we estimate an *RCL* (rounded to 2 significant figures) of 12 ug/kg (*y* value) for TCE.

Linear Fit Between SPLP leachate and Soil data. Will also plot Prediction Limit Estimates on the Line Fit. Ref: ***Weisberg*** *[1980].*

Sample Designation	C_s Total TCE ug/kg	C_l SPLP TCE ug/l
B-29	2.0	0.1
B-37(1-3)	8.0	1.0
B-37(4-6)	15.0	1.7
B-38(5-7)	20.0	1.9
B-37(5-7)	50.0	5.0
B-26(10-11)	72.0	2.0
B-23(6-7)	91.0	2.0
B-41	94.0	2.5
B-45	130.0	60.0
B-45(8-10)	170.0	57.0
B-18	175.0	35.0



α =	0.05			
Prediction Interval for RCL(%) =	95	95% of all possible samples would yield interval limits capturing RCL.		
		RCL⁻	RCL	RCL⁺
PAL (ug/l) =	0.5	1.7	11.5	79.4

There are very important subtleties to our analysis. The least-square predictor line is shown as solid red line. Microsoft Excel will automatically determine this line (and its equation) after we: 1.) Plot the data, 2.) "Click" on one of the plotted points, 3.) Under "*Chart*," click "*Add Trendline,*" 4.) Select "*Power*" type, and 5.) Under "*Options*," check the boxes for "*Display equation…*" and "*Display R-squared…*"

For the TCE data example, the equation for the predictor line ($y = 17.83\ x^{0.6354}$) is shown with the associated coefficient of determination R^2. Microsoft Excel calculates the R^2 using a log-transformed regression. For most of our purposes, this R^2 should be satisfactory in judging how well the line fits the data. An $R^2 > 0.6$ shows strong correlation when the number of data pairs is 10 or more.

The Excel routine, however, will *not* provide the statistical confidence interval about the fitted values. In the above plot, the broken curves are the upper (RCL+) and lower (RCL–) limits showing the 95% confidence interval for the fitted value C_s (or y) for a given C_l (or x). The confidence interval limits creates a band about the predictor line [*e.g. Weisberg*, 1980; *Larsen and Marx*, 2001]. Note that the 80% confidence band would be *trimmer* compared to higher confidence bands. If you see some curvature in both the RCL+ and RCL– lines, you are not seeing things! Both the RCL+ and RCL– lines curve away from the central *x* value. In the example, the central *x* value is the

geometric mean (GM of 3.7 ug/l) of the C_l (or x) data. The curvature is because the statistical analysis of the data suggests less confidence for predicting C_s (or y) *away* from where the C_l data are centered. A statistically more robust RCL would be estimated when the SPLP leachate C_l data is centered nearer the *PAL* level of interest.

In our example, TCE's *PAL* of 0.5 ug/l is "away from" the data GM of 3.7 ug/l (by more than a log unit). Our confidence on an RCL of 12 ug/kg may *not* be optimal because of the paucity of data in the low-end leachate range (0.2 to 0.9 ug/l). The 95% prediction interval shows that the RCL can be as low as 1.7 ug/kg (or lower than the generic RCL of 3.7 ug/kg shown in p. 16 of http://www.dnr.state.wi.us/org/aw/rr/archives/pubs/RR682.pdf). Hence by considering this statistical confidence, we determine that the soil-to-groundwater RCL at this site must NOT exceed 12 ug/kg.

It is important that both K and n (from our analysis) must be positive numbers, or else the results do not have any meaning in the physical sense. (If n is negative, for instance, we would predict - incorrectly - that a soil contaminant concentration smaller than our PAL-based RCL would be leaching at ES.) When least-square analysis is performed on the log-transformed data ($\ln(C_s)$ and $\ln(C_l)$), the result may yield a negative n, so the simple Excel routine (as described here) will not always work. If the resulting n is negative, then there may not be adequate data or the adsorption phenomenon may be better described by the Langmuir isotherm or another type of empirical relation [*Bohn et al*., 1979]. We provide at the end of this addendum a separate example to illustrate how to determine K and n when the simple Excel analysis yields a negative parameter. Briefly, the technique iteratively searches for the parameters without necessarily requiring any logarithmic transformation of the data [*e.g. Press et al*., 1986].

Details on Alternative Step 7 - Using the C_s / C_l (or "Y/X") Ratio Statistics. We illustrate the analysis that was described in section "B" (*Use of Leaching Test Data*) of the SPLP guidance (PUBL RR-523-03, p. 8). (Again, we will ignore the data pairs when the dry-soil contaminant concentrations are over the cut off that we determined earlier.) We can then compare our RCL estimate (12 ug/kg) from Steps 5 & 6 with the result of this alternative. Section "B" calls for 3 things: 1.) Take the ratios "Y/X"; 2.) Estimate the 95% lower confidence limit of the mean (*LCL*) of the ratios; and 3.) *RCL = LCL x PAL*.

1.)

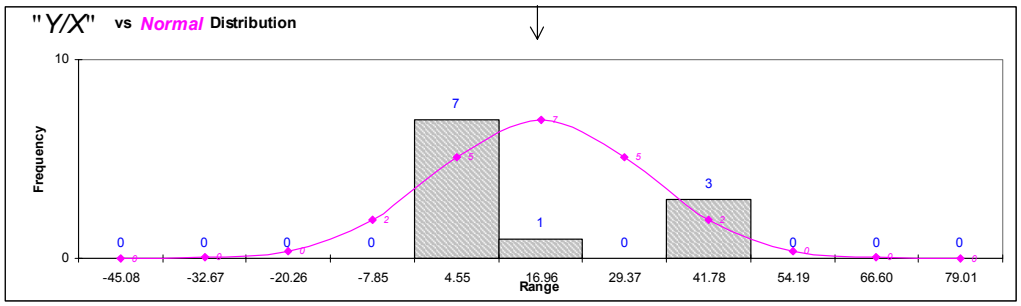
Sample	Total TCE (ug/kg)	SPLP TCE (ug/l)	"Y / X"	ln (Y/X)
B-29	2.0	0.1	20.00	2.995732
B-37(1-3)	8.0	1.0	8.00	2.079442
B-37(4-6)	15.0	1.7	8.82	2.177422
B-38(5-7)	20.0	1.9	10.53	2.353878
B-37(5-7)	50.0	5.0	10.00	2.302585
B-26(10-11)	72.0	2.0	36.00	3.583519
B-23(6-7)	91.0	2.0	45.50	3.817712
B-41	94.0	2.5	37.60	3.627004
B-45	130.0	60.0	2.17	0.773190
B-45(8-10)	170.0	57.0	2.98	1.092747
B-18	175.0	35.0	5.00	1.609438

"Y/X" Sorted

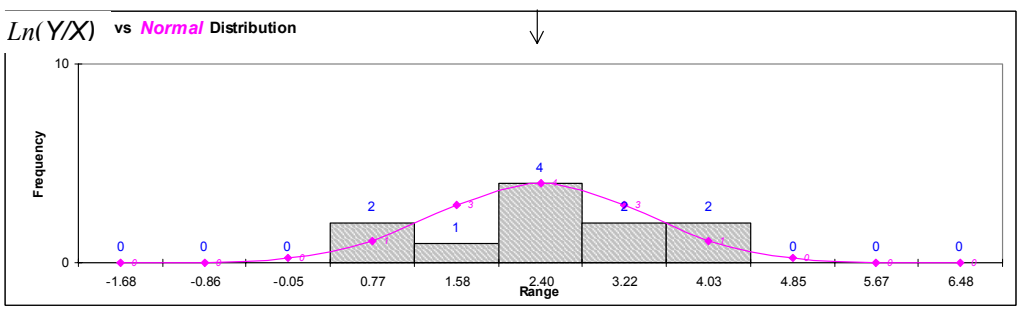
"Y / X"	ln (Y/X)
2.17	0.773190
2.98	1.092747
5.00	1.609438
8.00	2.079442
8.82	2.177422
10.00	2.302585
10.53	2.353878
20.00	2.995732
36.00	3.583519
37.60	3.627004
45.50	3.817712

2.) *LCL Estimate Depends on Data Distribution.* The typical hang up in doing the *Step 7* procedures is in *not* knowing whether the ratios are distributed normally or lognormally (*i.e.* "lognormal" is when it is the logarithms of the ratios that are normally distributed). The type of distribution would dictate how the *LCL* is estimated. The *LCL* estimate will involve the more familiar *t* statistic (for normally-distributed data) or the less familiar *H* statistic (for lognormally-distributed data).

The simplest way to show which of the 2 distributions may be more appropriate is by comparing the histograms of the "Y/X" and the Ln(Y/X) sets as shown below:



The "Y/X" values appear to be bimodal. The mean value (16.96) may be too high for most of the data. (*Option:* The 3 highest values may be deleted to see if the remaining data may be more normally distributed, and then a new mean is estimated.)



The transformed ratios, Ln (Y/X), seemed to be more normally distributed. The geometric mean value ($e^{2.43} = 11$) would better describe the central tendency of the ratios.

With the above histograms, we show that the ratios are more likely to be lognormally distributed. We can then proceed with calculating the *LCL*. (When there is sufficient data, another option at this point is to delete the high *untransformed* ratios, which may help show that the remaining ratios are normally distributed, and then the *LCL* is determined from the arithmetic mean of the remaining ratios and *t* statistics.)

Determining the Distribution Using the Nonparametric Shapiro-Wilk W Test Another option that can help decide if the ratios are normally or lognormally distributed is through the W statistic and the nonparametric Shapiro-Wilk W test [*Gilbert*, 1987]. The W test can be used to *reject* the (null or default) hypothesis that a set of data is normally distributed. However, when the W test fails to reject the default hypothesis, it only implies that the normal distribution is one type of potentially many possible distributions that can describe the data. For instance, we can use the W test on a *set* of data Ÿ, and also on *ln*(Ÿ). Doing so may result in the W test failing to reject the normal distribution hypothesis on both Ÿ and *ln*(Ÿ), implying that we can not favor the normal over the lognormal (*and* vice versa) for our Ÿ data. We can: 1.) Collect additional data to help resolve the true shape of the distribution, or 2.) When there are sufficient (15 or more)

data in the subset, increase the α level of the W test (as a larger α increases the rejection area of the null hypothesis). When we do the latter (*i.e.* increase α), the W test may prove useful in rejecting one of the assumed distributions. For illustration, let's do this latter procedure on the subset series of "Y/X" and ln(Y/X) given above (even though we only have 11 points). The W statistics are 0.8263 and 0.9434 for the untransformed and log-transformed ratios, respectively. The results of the "comparative" W test procedure (pitting the normal-distribution hypothesis against the lognormal case) are shown below.

A Program by Resty Pelayo, WDNR, 608/267-3539

W Test for Normal Distribution (from Gilbert, Statistical Methods for ..., p. 158-160); H_o: the data has a normal distribution. W test is for the rejection of H_o!

Instructions: 1.) Input raw (unsorted) data Under x_i column; 2.) Click "Sort xi" button; 3.) Compare the W to W_p in Table A7 in p. 261, Gilbert;
4) If W < Wp: Reject normality at level p (easier to reject at larger p); 5) If W > W_p: Can NOT reject - Pass test (i.e., OK to assume normality at p significance level - smaller p is better)

INPUT										
Soil Concentration								Normal	Lognormal	
Raw Data	Clear Entries		Sort xi					untransformed	transformed	
x_i	$y_i = \ln(x_i)$		Ordered x_i	Ordered y_i				x_i	y_i	
20	2.9957	1	2.166666667	0.7732		n: # of data	11	11	11	<== If not all equal, click on "Sort xi" button.
8	2.0794	2	2.98245614	1.0927		Mean		16.9635	2.4012	
8.823529412	2.1774	3	5	1.6094		Std.Dev.		15.511	1.021	
10.52631579	2.3539	4	8	2.0794					10.4154	
10	2.3026	5	8.823529412	2.1774		d: sum of squared deviations		2405.8780	10.4154	
36	3.5835	6	10	2.3026		k: median locator	5			
45.5	3.8177	7	10.52631579	2.3539				$W = [\Sigma\, a_i\,(x_{n-i+1} - x_i)]^2 / d$	$W = [\Sigma\, a_i\,(y_{n-i+1} - y_i)]^2 / d$	
37.6	3.6270	8	20	2.9957				Normal	Lognormal	
2.166666667	0.7732	9	36	3.5835		W: test statistic -->		**0.8263**	**0.9434**	<== Smaller W: Reject normality hypothesis.
2.98245614	1.0927	10	37.6	3.6270		p	W_p	$W < W_p$?	$W < W_p$?	Normal Hypothesis / Lognormal Hypothesis
5	1.6094	11	45.5	3.8177		0.01	0.792	No - Can not reject Normality	No - Can not reject Lognormality	Fail to reject at level 0.01 / Fail to reject at level 0.01
		12				0.02	0.817	No - Can not reject Normality	No - Can not reject Lognormality	Fail to reject at level 0.02 / Fail to reject at level 0.02
		13			--->	0.05	0.85	Yes - Reject Normality	No - Can not reject Lognormality	Reject at level 0.05 / Fail to reject at level 0.05
		14				0.10	0.876	Yes - Reject Normality	No - Can not reject Lognormality	Reject at level 0.1 / Fail to reject at level 0.1
		15				0.50	0.94	Yes - Reject Normality	No - Can not reject Lognormality	Reject at level 0.5 / Fail to reject at level 0.5
		16						3	0	
		17						0.05	FALSE	
		18								
		19				**Conclusion about Normal Distribution:**		**REJECT normal distribution hypothesis at level 0.05 test.**		
		20								
		21				**Conclusion on comparing the 2 distributions: Easier to REJECT NORMAL distribution hypothesis.**				
		22								

Since we can reject the normal distribution hypothesis (but *fail to reject* the lognormal hypothesis) when the level α for the W test is increased to 0.05 or more, we have likewise shown that the ratios are probably better described as lognormally distributed.

Determine LCL. Our frequency plot and the W test support the ratios' lognormality. We can determine the *LCL* (or the one-tailed lower 95% confidence limit of the mean) by using *Gilbert's* [1987] equation 13.14 and *Gilbert's* [1987] Table A13 (for the *H* statistic):

$$LCL = e^{\left(y_bar + \frac{s_y^2}{2} + \frac{s_y H_{0.05}}{\sqrt{n-1}} \right)} \qquad \textbf{(5)}$$

y_bar=	2.401				
s_y =	1.021		$LL_{0.05}$ =	10.818	---> Eqtn 13.14
n =	11	log-normal mean =		17.416	---> Eqtn 13.3
$H_{0.95}$ =	3.05053		$UL_{0.95}$ =	49.720	---> Eqtn 13.13
$H_{0.05}$ =	-1.6755				

$$LCL = 11\ (\textit{rounded to 2 significant figures}) = LCL \text{ of } \frac{C_s}{C_l}$$

3.) The 3rd part is the easiest. Using equation (**4**), we get:

$$SSRCL(ug/kg) = \left(\text{LCL of } \frac{C_s}{C_l} \right) PAL = (11)\ 0.5 = 5.5$$

The result following alternative *Step 7* is an RCL of 5.5 ug/kg that falls between the lower 95% (RCL–) prediction limit (of 1.7 ug/kg) and prediction line (12 ug/kg) from *Steps 5 and 6*. From these results, we can draw the conclusion that the RCL for the site should be about 5.5 ug/kg and *must not* exceed 12 ug/kg.

CONCLUSION

Empirical estimates for a site-specific RCL are possible *via* analysis of SPLP data collected at the site. For the analysis to be meaningful, we have to make sure that we look at the correct spectrum of the data; *i.e.*, rather than rely on the high soil contaminant concentration data, we rely more on the low-concentration region in determining *PAL*-based RCLs. The design for collecting SPLP data should have this important consideration for a statistically robust estimate of site-specific RCLs. We described 2 simplified procedures on how to analyze SPLP data. Both procedures should converge on similar RCL estimates, which will indicate how robust the estimated RCL is.

For additional information, contact Resty Pelayo at (608) 267-3539.

REFERENCES CITED

Bohn, H., B. McNeal and G. O'Connor, 1979, *Soil Chemistry*, John Wiley & Sons, New York, 329 p.

Domenico, P. A. and F. W. Schwartz, 1990, *Physical and Chemical Hydrogeology*, John Wiley & Sons, New York, 824 p.

Gilbert, R. O., 1987, *Statistical Methods for Environmental Pollution Monitoring*, Van Nostrand Reinhold, New York, 320 p.

Larsen, R. J. and M. L. Marx, 2001, *An Introduction to Mathematical Statistics and Its Application*, Prentice-Hall, Inc., Upper Saddle River, NJ, 790 p.

Press, W. H., B.P. Flannery, S.A. Teukolsky and W.T. Vetterling, 1986, *Numerical Recipes, The Art of Scientific Computing*, Cambridge University Press, New York, 818 p.

Roy, W. R., I. G. Drapac, S. F. J. Chou, and R. A. Griffin, 1992, *Batch-type Procedures forEstimating Soil Adsorption of Chemical*s. EPA/530/SW-87/006-F, OSWER, U.S. Environmental Protection Agency, Washington, D.C., 100 p.

Weisberg, S., 1980, *Applied Linear Regression*, John Wiley & Sons, New York, 283 p.

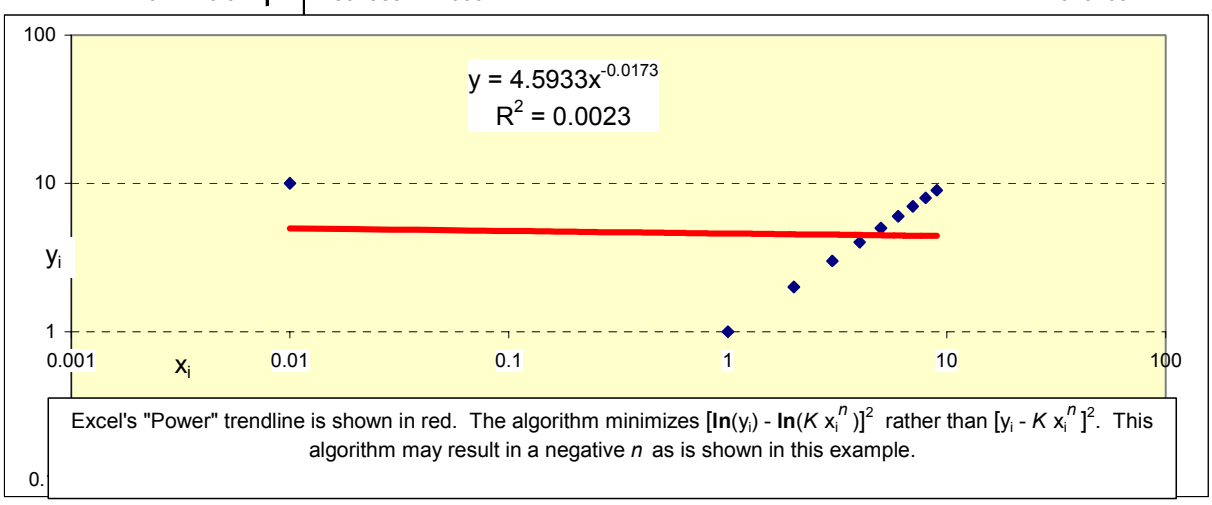
Zheng, C and G. D. Bennett, 2002, *Applied Contaminant Transport Modeling, 2nd Ed.*, John Wiley & Sons, New York, 621 p.

POSTSCRIPT: EXCEL'S "POWER" ROUTINE DOES *NOT* ALWAYS WORK

We provide another example data set to illustrate the general weighted least squares procedure [*e.g. Press et al.*, 1986] when fitting y data to an empirical function of x.

When the empirical Freundlich isotherm equation: $C_s = K\, C_l^{\,n}$ is log-transformed, we have the linear relation: $\ln(C_s) = n \ln(C_l) + \ln(K)$. The parameters K and n can be estimated when the linear least-square procedure is applied to the log-transformed SPLP data $\left(\ln(C_s)\right.$ and $\ln(C_l)$ as y and x, respectively$\left.\right)$. In fact, this is exactly the algorithm that Excel's *Power* trendline utilizes. However, one disadvantage in this algorithm is that it can not constrain that the slope n be positive (as is required in the Freundlich isotherm equation). The algorithm may yield an unusable negative parameter n as we show in the illustrated example data set below.

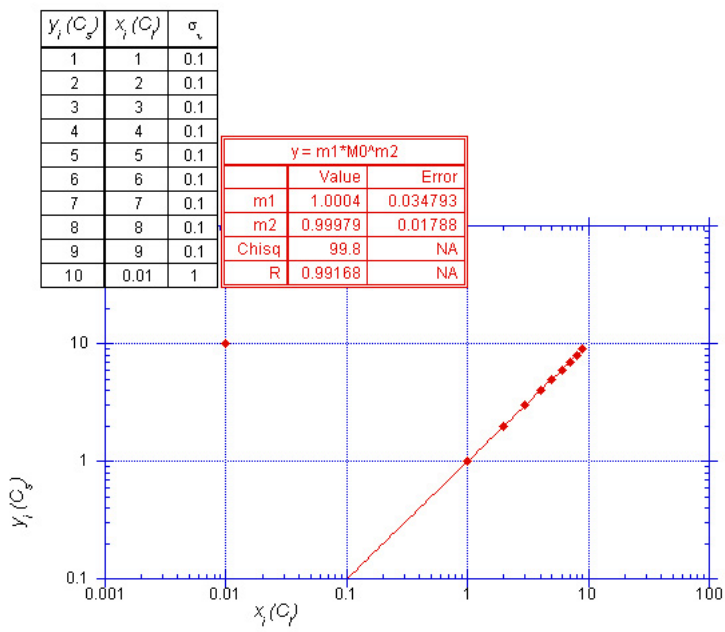
	INPUT DATA		OUTPUT			
Count	y_i (C_s)	x_i (C_l)	$\ln(y_i)$	$\ln(x_i)$		Pred. $y_i = K x_i^n$
1	1	1	0	0	slope: -0.017265593	4.593275
2	2	2	0.693147	0.693147	intercept: 1.524593272	4.538632
3	3	3	1.098612	1.098612	R^2: 0.002270772	4.50697
4	4	4	1.386294	1.386294	$\ln(y_i)$ = slope * $\ln(x_i)$ + intercept	4.484639
5	5	5	1.609438	1.609438		4.467395
6	6	6	1.791759	1.791759	K = exp(intercept): 4.593274975	4.453354
7	7	7	1.94591	1.94591	n = slope: -0.017265593	4.441517
8	8	8	2.079442	2.079442		4.431289
9	9	9	2.197225	2.197225	$y = K x^n$	4.422287
10	10	0.01	2.302585	-4.60517		4.973403



When we get a negative parameter using Excel's "Power" trendline (like getting $n = -0.0173$, above), a more appropriate procedure to analyze the data may be the general weighted least squares procedure [*Press et al.*, 1986]. The parameters n and K are determined by reducing the "goodness-of-fit" χ^2 error defined as $\chi^2 = \sum [(y_i - K\, x_i^{\,n})/\sigma_i]^2$ without resorting to logarithmic transformation of the data. The initial step is to assign measurement error σ_i on each data pair. A larger error σ_i will weigh the data pair down relative to the other measurements, so a data pair with smaller σ_i indicates how relatively

more confident we are about it compared to a data pair with a larger σ_i. An "outlier" can be assigned a larger σ_i. so that it does not affect the rest of the more coherent data.

While it is possible to program a script in Excel that can automate the task of determining *K* and *n* which will yield a minimum χ^2 (using Excel's "*Solver*," for instance), it may be easier to use other commercially available *scientific* spreadsheets. The result from one such scientific spreadsheet (Synergy Software's Kaleidagraph, see http://www.synergy.com) is shown in the figure below.



The curve-fitting equation used above was: $y = m1 \cdot x^{m2}$; m1 and m2 were determined given the paired set of observations (y_i , x_i) and the errors (σ_i) associated to each pair. We get: K = m1 = 1.0004; and n = m2 = 0.99979. The goodness-of-fit "Chisq" error is: $\chi^2 = \Sigma\, [(y_i - K\, x_i^{\,n})/\sigma_i]^2$. The apparent large χ^2 error in the example is attributed to a single data point (x_i = 0.01, y_i = 10): $\chi^2 \sim [10 - 1.0004\ (0.01)^{0.99977}]^2 = 99.8$. When we weighed this data pair down by assigning it a larger σ_i (*i.e.* $\sigma_i = 1$) compared to the rest of the data ($\sigma_i = 0.1$), the curve-fitting routine focused on the other data points. A high R (correlation coefficient) is associated with the particular results. With *K* and *n* (both positive numbers), we can proceed to use equation (**3**) to determine the RCL.

[While it is true that we could have deleted the outlier from our analysis, we may not have this luxury when faced with real site data. The example provides a procedure to downplay some data points, but not necessarily ignoring them, especially since these points will still show up on the data plot.]