EPA/600/R-06/047
ERASC-013
May 2006

ESTIMATION OF BIOTA SEDIMENT ACCUMULATION FACTOR (BSAF) FROM PAIRED OBSERVATIONS OF CHEMICAL CONCENTRATIONS IN BIOTA AND SEDIMENT

by

Lawrence Burkhard
U.S. Environmental Protection Agency
Office of Research and Development
National Health and Environmental Effects Research Laboratory
Mid-Continent Ecology Division
Duluth, Minnesota

Ecological Risk Assessment Support Center
Office of Research and Development
U.S. Environmental Protection Agency
Cincinnati, OH

NOTICE

This report is an external draft for review purposes only and does not constitute Agency policy. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

TABLE OF CONTENTS

AUTHORS, CONTRIBUTORS AND REVIEWERS

AUTHOR

Lawrence Burkhard
U.S. Environmental Protection Agency
Office of Research and Development
National Health and Environmental Effects Research Laboratory
Mid-Continent Ecology Division
Duluth, MN 55804

REVIEWERS

Keith Sappington
U.S. Environmental Protection Agency
Office of Research and Development
National Center for Environmental Assessment
Washington, DC 20460

Steven Ferraro
U.S. Environmental Protection Agency
Office of Research and Development
National Health and Environmental Effects Research Laboratory
Western Ecology Division
Newport, OR 97365

ACKNOWLEDGMENTS

Programmatic review of the document was conducted by Dale Hoff of EPA Region 8, a Trichair of EPA's Ecological Risk Assessment Forum. Jeff Swartout of EPA's National Center for Environmental Assessment, ORD, kindly provided review of BSAF statistical approaches.

INTRODUCTION

In March 2004, the Ecological Risk Assessment Forum (ERAF) submitted a request to ORD's Ecological Risk Assessment Center (ERASC) relating to the estimation of Biota-Sediment Accumulation Factors (BSAFs) (Appendix). BSAF is a parameter describing bioaccumulation of sediment-associated organic compounds or metals into tissues of ecological receptors. The *Problem Statement* in the request was "What is the most appropriate method to estimate the BSAF from paired observations of concentrations in biota and sediment?" The *Expected Outcome* asked for answers to specific questions regarding the use of regression analysis for estimating BSAFs for nonionic organic compounds. The specific questions are addressed in the latter portion of this document. A statement on the most appropriate method to estimate the BSAF is provided below. This document is focused solely on the determination of BSAFs for nonionic organic chemicals and is primarily applicable to fish and high level shellfish, e.g., crabs. The determination of BSAFs for metals is not discussed.

RECOMMENDATIONS

There are two methods for determining the BSAF from paired observations: 1) a regression approach, whereby the BSAF is estimated by determining the slope of the C_{soc}-C_{ℓ} line [C_{soc} is the concentration of chemical in the sediment on an organic carbon basis (µg/kg organic carbon) and C_{ℓ} is the concentration of chemical in the organism on a lipid basis (µg/kg lipid)], and 2) an averaging approach, whereby the BSAF is estimated by averaging the BSAFs from the paired observations across the site. Both approaches use the same data. The second approach, however, is generally the more appropriate method for estimating the BSAF because regression analysis has these four limitations:

1) Regression analysis, whether model I (simple linear regression) or model II (geometric mean regression, major axis regression, Bartlett's three-group method, or Kendall's robust line-fit method (Sokal and Rohlf, 1995)), requires meeting parametric assumptions about the relationship between the X and Y variables.

2) Regression analysis, in order to be useful, requires a range of values in the X and Y variables.

3) When large ranges exist in the C_{soc}-C_{ℓ} values (e.g., C_{soc} spans two orders of magnitude), weighting of the data in the regression analysis and/or transformation of the data might be required for proper analysis.

4) Although regression analysis can be done on data sets with limited numbers of C_{soc}-C_{ℓ} pairs, determining the slope of the line fitting limited numbers of pairs can lead to highly uncertain slopes.

In contrast, the averaging approach (estimating the BSAF by averaging the BSAFs from each C_{soc}-C_{ℓ} pair) requires none of these conditions or assumptions. Further, unlike the regression approach, the averaging approach can be performed with limited data.

Both the regression and averaging approaches require similar conditions (e.g., food web structure, sediment/water column concentration quotients, chemical bioavailability, and diets of the organisms) for each C_{soc}-C_{ℓ} pair. (This can be problematic for Superfund and other sites that have highly heterogenous conditions.) Additionally, for both approaches, accuracy and precision of the calculated BSAFs are a function of the sample size, i.e., the number of the C_{soc}-C_{ℓ} pairs.

With the regression and averaging approaches, each C_{soc}-C_{ℓ} pair is location specific and each pair incorporates all of the conditions existing at the location. In order to use either approach, the conditions must be the similar across all locations. Mixing of C_{soc}-C_{ℓ} paired observations with different underlying conditions is not recommended and will, in all likelihood, result in BSAFs with poor predictive accuracy.

With the averaging approach, the distribution of the individual BSAFs (determined from each C_{soc}-C_{ℓ} pair) can be evaluated very easily; this evaluation is commonly done in statistical analysis of data. Knowing the underlying distribution of the BSAFs allows the selection of the most appropriate (unbiased) averaging technique. Further, with the individual BSAFs (C_{soc}-C_{ℓ} pairs), the homoscedasticity (equality) of the variances across the individual BSAFs can be assessed. In cases where the variances are heteroscedastic (unequal), an appropriate weighted averaging technique would be used, and in general, the weights would be the reciprocal of the variances for the individual BSAFs. The averaging approach can also be easily implemented with other weighting considerations such as portions of the site represented by individual BSAFs, e.g., some BSAFs might be reflective of three quarters of the site while the remaining BSAFs are reflective of the other quarter of the site. The averaging approach also provides the information on the final BSAF (grand mean) distribution and variance which are required for one and two stage Monte Carlo uncertainty analyses.

There is great value in plotting the C_{ℓ} against C_{soc}; BSAFs against C_{soc}; and C_{ℓ}, C_{soc}, and BSAFs against geographical information. These plots should be done and evaluated for trends in the data! They may provide key insights and understanding of the complexities existing at the site of interest. The importance of resolving discrepancies within the data can not be overstated (e.g., Why are some BSAFs so different? Are there trends or dependencies upon concentrations of chemicals in sediment or with geographical location within the site? Why don't the C_{soc}-C_{ℓ} pairs form a linear relationship?) Spending time and resources resolving these discrepancies will be well worth the effort since the uncertainties associated with remediation decisions will be

smaller. Additionally, any discrepancies in the data at this level will be translated into higher and more complex analyses since these analyses use this information.

The following sections provide a description of the BSAF along with its underlying assumptions, a discussion on how to measure a useful BSAF, a discussion on the basis of the regression approach, and answers to specific questions related to regression analysis.

DEFINITION OF BSAF

The BSAF is defined (Ankley et al., 1992) as

$$BSAF = \frac{C_o/f_\ell}{C_s/f_{soc}} \quad \textbf{(1)}$$

where C_o is the chemical concentration in the organism (µg/kg wet weight), f_ℓ is the lipid fraction of the organism (g lipid/g wet weight), C_s is the chemical concentration in surficial sediment (µg/kg dry weight) and f_{soc} is the fraction of the sediments as organic carbon (g organic carbon/g dry weight). In general, BSAFs should be determined from spatially and temporally coordinated fish and surficial sediment samples under conditions in which recent loadings of the chemicals to ecosystem are relatively unchanged (Burkhard et al., 2003). The BSAF definition does **not** invoke or include the assumption of equilibrium conditions for the chemical between the organism and sediment (Ankley et al., 1992; Thomann et al., 1992). As shown by Thomann et al. (1992), BSAFs are appropriate for describing bioaccumulation of sediment contaminants in aquatic food webs with non-equilibrium conditions between both the sediment and fish, and sediment and its overlying water. Equilibrium is regarded as a reference condition for describing degrees of disequilibrium, and thus, is not a requirement for measurement, prediction, or application of BSAFs.

With specific reference to benthic invertebrates, numerous investigators (Lake et al., 1984; McElroy and Means, 1988; Bierman, 1990; Lake et al., 1990; Ferraro et al., 1990) have invoked two assumptions regarding BSAFs: 1) equilibrium conditions and 2) no metabolism of the chemical. These assumptions when combined with EqP (equilibrium partitioning) theory (DiToro et al., 1991), leads to the conclusion that the BSAF, for these specific conditions, is equal to the partitioning relationship of the chemical between organic carbon in the sediment and lipids of the organism. Depending upon the affinities of the nonpolar organic chemical for lipid and sediment organic carbon, the BSAF, under these specific conditions, should be in the range of 1 to 2 (McFarland and Clarke, 1986). For aquatic organisms tightly connected to the sediments like oligochaetes and other benthic invertebrates, experimental measurements (Lake et al., 1990; Tracy and Hansen, 1996) are generally consistent with the theoretical value, i.e., in the range of 1 to 2.

There are solid mechanistic reasons why fish should not be in equilibrium with their sediments (Thomann et al., 1992). For fish, BSAFs incorporate wide ranges of influences including biomagnification due to the trophic level of the fish; sediment-water column chemical disequilibrium; the diet of the fish and its underlying food web; the fish's home range, and chemical metabolism within the fish and its food web (Burkhard et al., 2003). Suggestions that BSAFs for fish should be in the range of 1 to 2 by combining the definition of the BSAF with the assumptions of equilibrium conditions and no metabolism are incorrect (Wong et al., 2001). As explained above, measured BSAFs above or below 1 to 2 are entirely reasonable for fish (Burkhard et al., 2003). BSAFs outside this range for fish do not violate the general definition of

BSAFs nor invalidate the usefulness of BSAFs in predicting chemical residues in fish for sediment contaminants (Burkhard et al., 2004).

MEASURING USEFUL C_{soc}-C_ℓ PAIRS FOR CALCULATION OF BSAFs

Probably the most important factor in measuring a BSAF with predictive power is the requirement that the sediment samples analyzed be reflective of the immediate home range of the fish. Depending upon the site, the degree of difficulty in defining the immediate home range of the organism can vary widely. In situations where the movement of the organisms is confined by the geography of the site, e.g., dams or falls, the home range of the organisms can probably be defined fairly easily. When required, home ranges can be determined by tagging/recapture, radio-telemetry, and/or ultrasonic telemetry studies at the site of interest. Estimates of home ranges for freshwater fishes can be determined using the allometric relationship (Minns, 1995):

$$\ln H = -2.91 + 3.14\ HAB + 1.65 \ln L \quad \text{or} \quad \ln H = 3.33 + 2.98\ HAB + 0.58 \ln W$$

where H is the home range size (m^2), HAB is 0 for rivers and 1 for lakes, W is body weight (g), and L is body length (mm). For freshwater invertebrates (crabs), marine and estuarine ecosystems, allometric relationships for home range have not been reported.

Having a good understanding of the immediate home range of the species is important. Organisms with smaller home ranges will, in all likelihood, be more representative of the study site than those with large home ranges that extend way beyond the study site. Just because a fish (or other aquatic organism) is caught at a sampling location, one can not infer that the chemical residue in the fish is due to the chemicals residing at the study site. Knowledge of the fish's home range is the only way that one can establish the connection of the fish to the sampling location. It is strongly recommended that local fisheries experts be consulted during the

sampling design phase of the field study to help in determining the immediate home range and trophic level of the organisms at the site; local knowledge will be extremely helpful. Although the above allometric relationship is available for estimating home ranges, one shouldn't necessarily assume that the "calculated" and "actual" immediate home ranges for the organisms are the same; one will still need to do the leg work of establishing as best as one can the immediate home ranges for the organisms at the site.

Once the home range of the species of interest is established, sediment samples reflective of the species home range need to be collected. It is important that the sediment samples collected be representative of the sediments to which the organisms are exposed and not a homogenized sediment core representing the entire bed of contaminated sediment. For most organisms, the surficial sediments are most reflective of the organism's immediate exposure history, and generally, smaller depths of the surficial layer, e.g., 0 to 2 cm, are preferred over larger depths, e.g., 0 to 30 cm. For deeper burrowing organisms such as some clams and polychaetes, slightly larger surficial depths, e.g., 0 to 5 cm, might be more appropriate of their recent exposure history.

Beyond establishing the home range of the organism and the appropriate sediment samples, the collection and analysis of adequate numbers of organisms and sediment samples is required for deriving unbiased estimates of the mean concentrations of chemicals with known variances. This document will not address the subject of sample collection, compositing, and analysis. With unbiased estimates of the mean concentrations, the BSAF for the specific site can be calculated using Equation 1.

In any study design, it is important that biota samples be collected and composited in size or age classes. For fish, dietary composition changes substantially with size and age, and these changes will result in differences in BSAFs among size and age classes. For forage fish, common classes are young-of-the-year, juveniles, and adults, and for piscivorous fishes, common classes are year classes, e.g., 2, 3, 6, and 10 years old. Mixing of fishes of different size/age classes is not recommended because of the increased variance for the average chemical residue in the organisms.

Biota samples for chemical analysis should never be composited by mixing different fish species. Different fish species have different life histories and diets. BSAFs derived from composite samples composed of different species will be highly biased by the individual species. Further, resolving what the potential biases are for an individual species would require the collection and analysis of that species.

When a C_{soc}-C_{ℓ} pair (or BSAF) is measured for a specific chemical, the measured value incorporates all conditions and parameters existing at the location of interest. The major conditions and parameters incorporated into the C_{soc}-C_{ℓ} pair (or BSAF) are 1) the distribution of the chemical between the sediment and water column, 2) the relationship of the food web to water and sediment, and 3) the length of the food web (or trophic level of the organism).

CALCULATION OF BSAFs

The BSAF is calculated from four measured variables (see equation 1, repeated below): concentration of the chemical in the organism on a wet weight basis (C_o), the lipid content of the wet tissue (f_{ℓ}), the concentration of the chemical in the sediment on a dry weight basis (C_s), and the organic carbon content of the dry sediment (f_{soc}).

$$BSAF = \frac{C_o/f_\ell}{C_s/f_{soc}} \quad \textbf{(1)}$$

A C_{soc}-C_ℓ pair will, in many cases, be composed of multiple composite tissue samples and multiple sediment samples (spanning the immediate home range of the organisms) for a sampling location. In order to determine the BSAF for the C_{soc}-C_ℓ pair, average concentrations in the tissue and sediment need to be determined; the numerator and denominator of Equation 1. The lipid normalized concentration of the chemical in each tissue sample should be determined and then, these values should be averaged to determine the average chemical concentration for the organisms. If the tissue samples have different numbers of organisms in each composite, e.g., three fishes in one sample and five fishes in the second sample, a weight average concentration should be determined. For normally distributed residues and the two sample fish example, the weighted average concentration equals:

$$C_{\ell-avg} = \sum(w_i \times C_{\ell-i}) / \sum w_i = (3 \times C_{\ell-one} + 5 \times C_{\ell-two})/(3+5) \quad \textbf{(2)}$$

where w_i is the number of organisms in composite i, $C_{\ell\text{-}i}$ is the lipid normalized concentration of the chemical in composite i, and $C_{\ell\text{-avg}}$ is the weighted average lipid normalized concentration in the tissues. The standard deviation of a weighted average ($s_{C\ell\text{-avg}}$) equals

$$s_{C_{\ell-avg}} = \sqrt{\left(\sum w_i \times (C_{\ell-i} - C_{\ell-avg})^2\right) / \left(\sum w_i - 1\right)} \quad \textbf{(3)}$$

For log-normally distributed residues in the fish, the weighting would be done on the log transformed data. Sediment samples would be treated similarly; normalizing for organic carbon and then, calculating the average concentration of the chemical in the sediments.

The BSAF for the C_{soc}-C_{ℓ} pair would then be determined by dividing $C_{\ell\text{-avg}}$ by $C_{soc\text{-avg}}$. The variance for the BSAF can be estimated using the equation (Mood et al., 1974):

$$s_{BSAF} = \frac{1}{C_{soc-avg}} \sqrt{(s_{C_{\ell-avg}})^2 + BSAF^2(s_{C_{soc-avg}})^2 - 2rs_{C_{\ell-avg}} s_{C_{soc-avg}} BSAF} \quad \textbf{(4)}$$

where s_{BSAF}, $s_{Csoc\text{-}avg}$, and $s_{C\ell\text{-}avg}$ are the standard deviations for the BSAF, $C_{soc\text{-}avg}$, and $C_{\ell\text{-}avg}$, respectively; and r is the correlation coefficient between $C_{soc\text{-}avg}$ and $C_{\ell\text{-}avg}$.

For each C_{soc}-C_{ℓ} pair, a BSAF is determined. As discussed previously, the average BSAF would subsequently be determined from the individual BSAFs using the most appropriate (unbiased) averaging technique based upon the underlying distribution of the BSAFs.

BASIS FOR BSAF REGRESSION APPROACH

Equation 1 can be rearranged:

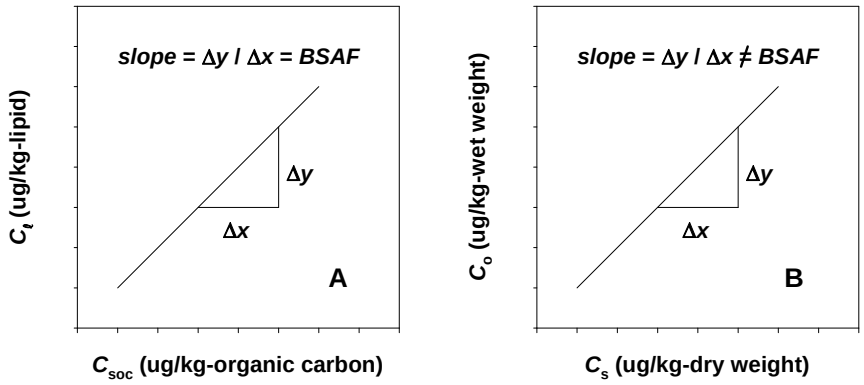
$$C_o/f_{\ell} = BSAF \times C_s/f_{soc} \quad \textbf{(5)}$$

By substitution, equation 5 can be expressed as:

$$C_{\ell} = BSAF \times C_{soc} \quad \textbf{(6)}$$

where C_{soc} is the concentration of chemical in the sediment on an organic carbon basis (µg/kg organic carbon) and C_{ℓ} is the concentration of chemical in the organism on a lipid basis (µg/kg lipid).

Plotting of C_{soc} against C_{ℓ} results in the following illustrative plot (Graph A), where the slope of the line is the BSAF. However, the slope of C_s plotted against C_o (Graph B) is not the BSAF because these two measures of chemical concentrations are **not** organic carbon and lipid normalized. Use of the regression approach to derive the BSAF incorporates an implicit assumption above and beyond those required for measuring a BSAF at a specific location. The implicit assumption of the regression approach is that all C_{soc}-C_{ℓ} pairs must have or incorporate the same underlying ecological conditions and parameters.



For a Superfund site, it is common to collect samples across the site with a number of different sampling locations. For example, consider a New England stream with a series of three dams, and assume that two-year-old carp and sediment are collected and analyzed in each reservoir. Further assume that enough fish and sediment were collected so that representative and unbiased mean concentrations were determined for each reservoir. Thus, three sets of paired carp-sediment observations would be determined, one for each of the three reservoirs.

These paired observations of C_{soc} and C_{ℓ} can be plotted (Graphs C & D). In Graph C, the pairs form a nearly linear relationship suggesting that the underlying conditions for the C_{soc}-C_{ℓ} pairs are consistent across the samples and thus allow estimation of the BSAF using the regression approach. In Graph D, the pairs form no easily defined linear relationship, and in this case, there is too little variability in the C_{soc}-C_{ℓ} pairs for the regression approach to be useful in estimating the BSAF. In Graph E, a situation where four sets of paired carp-sediment data were determined, three of the pairs form a nearly linear relationship, but one pair is different from the other pairs. Depending upon how one draws the line, either the triangle or square data in Graph E could be the different (or outlier) C_{soc}-C_{ℓ} pair. In this case, one or more of the C_{soc}-C_{ℓ} pairs have different underlying conditions, and thus, it would be inappropriate to estimate the BSAF using the regression approach.

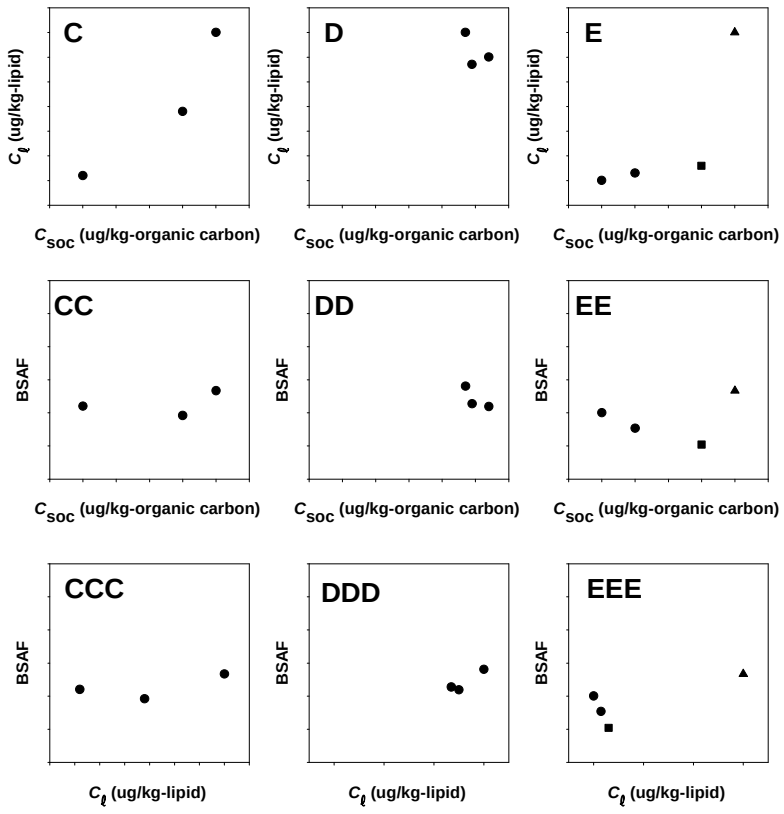
As discussed above, each carp-sediment pair is location specific and each pair incorporates all of the major conditions and parameters existing at the location. In order to use the regression approach with pairs of C_{soc}-C_{ℓ} observations, the major conditions and parameters must be the same for all locations. This requirement is the implicit assumption incorporated into the regression approach. Mixing of C_{soc}-C_{ℓ} paired observations with different conditions and parameters will result in C_{soc}-C_{ℓ} plots where the C_{soc}-C_{ℓ} pairs will form a non-linear relationship (e.g., possibly Graph E), and in all likelihood, a BSAF with poor predictive power.[1]

For the above examples, if the BSAF for each pair of C_{soc}-C_{ℓ} observations are plotted against C_{soc}, the following graphs are obtained (graphs CC, DD, and EE). The relationships among the C_{soc}-C_{ℓ} pairs in the above graphs remain in the graphs based upon the BSAFs;

[1]The mixing of C_{soc}-C_{ℓ} paired observations with different conditions and parameters is not recommended for the averaging approach as well. BSAFs with poor predictive power (i.e., accuracy) will, in all likelihood, result when different conditions and parameters exist across the individual C_{soc}-C_{ℓ} pairs used in the analysis.

compare Graphs C to CC, D to DD, and E to EE. In essence, by calculating the BSAF, one has mathematically removed the concentration dependence shown in Graphs C, D, and E. For further comparison purposes, the BSAF for each pair of C_{soc}-C_{ℓ} observations are also plotted against C_{ℓ} (graphs CCC, DDD, EEE).

The graphs, i.e., C, D, E, CC, DD, EE, CCC, DDD, and EEE, are some of the plots recommended for evaluating trends and underlying conditions associated with the C_{soc}-C_{ℓ} pairs. We recommend that these plots be completed prior to performing the final calculations for determining the site-specific BSAF. These plots will help in identifying sources of variation and error in the individual C_{soc}-C_{ℓ} pairs and BSAF values.



THE REGRESSION APPROACH

A key consideration in using the regression approach is to realize that both C_{soc} and C_{ℓ} are measured with error. With the simple linear regression least-squares technique, one variable (the *Y*'s) are measured with error while the other variable (the *X*'s) are fixed and have no error. Simple linear regression is referred to as model I regression analysis. When *X*'s and *Y*'s are both measured with error, one of a number of model II regression techniques will be more appropriate and unfortunately "the appropriate method depends on the nature of the data" (Sokal and Rohlf, 1995). Sokal and Rohlf (1995) provide an excellent discussion on model II regression and the techniques of geometric mean regression (also called reduced major axis, standard major axis, or relation d'allometrie), slope of the major axis, Bartlett's three-group method, and Kendall's robust line-fit method. Additionally, Sokal and Rohlf (1995) discuss the *Berkson case* of model II regression where model I regression is appropriate.

It is suggested that the determination of the slope of C_{soc}-C_{ℓ} pairs be performed using the geometric mean regression technique (Halfon, 1985; Sokal and Rohlf, 1995) because with this technique the slope of the regression is not dependent upon the scale of the *X*'s and *Y*'s used in the analysis. Additionally, Ricker (1973) has recommended that the geometric mean regression technique be used for determining functional relationships (i.e., slope) when "the variability is mostly natural ... in *X* and *Y*"; the case, I believe, when sediment samples representative of the organism's actual exposure history are collected.

For the geometric mean regression technique, the slope of geometric mean regression line is the geometric mean of the slopes of the following two linear regression least-squares lines:

$$y = a + b'' x \qquad \textbf{(7)}$$

and

$$x = c + dy \quad \textbf{(8)}$$

The slope of the geometric mean regression line is computed as the geometric mean of b'' and $1/d$:

$$b = (b'' / d)^{1/2} \quad \textbf{(9)}$$

The intercept a is computed as done in linear regression:

$$a = \hat{Y} - b\bar{X} \quad \textbf{(10)}$$

For further details on the geometric mean regression technique, the reader is referred to Halfon (1985) and Sokal and Rohlf (1995).

An Excel add-in function for geometric mean regression can be downloaded from the following URL.

http://www.uottawa.ca/academic/arts/geographie/lpcweb/newlook/data_and_downloads/download/sawsoft/modelii/modelii.htm

RESPONSES TO QUESTIONS RAISED IN EXPECTED OUTCOMES

Do I fit a straight line through the data?

Yes. If the C_{soc}-C_{ℓ} observations don't form a straight line, then one must figure out why data diverge from the linear relationship. Reasons for the C_{soc}-C_{ℓ} observations diverging from a straight line include (Note, there are many more causes than those listed):

- The organisms in different C_{soc}-C_ℓ pairs reside at different trophic levels in the food web.

- The organisms in different C_{soc}-C_ℓ pairs have dramatically different diets even though they reside at the same trophic level. For example, for one pair, the organisms might consume primarily zooplankton while for other pairs, the organisms might consume primarily benthic invertebrates.

- The bioavailability of the chemical in the contaminated sediment varies substantially across the C_{soc}-C_ℓ pairs.

- Across the sampling locations, inputs of the chemicals to the site differ substantially. For example, consider a harbor where organisms residing in the lower parts of the harbor are exposed to runoff and ground water seepage from an old industrial site while organisms residing in the upper parts of the harbor are not exposed this to discharge.

- Different populations of the same species. For example, in the Hudson River, there are resident and migratory striped bass fish populations, and chemical residues in the populations differ widely.

Do I plot my data on a log-log scale?

It is recommended that the data be plot in arithmetic-arithmetic scales because in arithmetic-arithmetic space, the slope of the line is the BSAF when C_{soc}-C_ℓ pairs are used. In general, the data, i.e., the C_{soc}-C_ℓ pairs, are assumed to be scaled arithmetically, and thus, should be plotted on arithmetic-arithmetic scales.

As a note of clarification, in log-log scales, the slope of the regression line (log C_ℓ regressed against log C_{soc}) is not the BSAF. See Equation 12, derived from the rearrangements of Equation 6 and then, Equation 11.

$$\log C_\ell = \log [C_{soc} \times BSAF] \quad \textbf{(11)}$$

$$\log C_{\ell} = slope \times \log C_{soc} + \log BSAF \quad \textbf{(12)}$$

Do I force the line through the origin?

Yes, when doing regression with arithmetic-arithmetic scales. (If one is performing the regression with log-log scales, the origin does not exist because the logarithm of zero is undefined. Thus, the line can not be forced through the origin.)

How do I handle non-detects?

I'm not sure of your definition of non-detects. I'll provide answers for both definitions: chemicals present at concentrations below the minimum detection limit (MDL) of the method and chemicals not detected at all, i.e., no response above instrumental noise. For the case where the chemical is present at concentrations below the MDL, use the uncensored value in the calculation; don't use the MDL value. For the case where the chemical is not detected at all, Superfund typically uses ½ of the MDL. However, as discussed below, there are approaches for working with data below the MDL and when the chemical is not detected at all. Calculation of BSAFs using arbitrarily ½ of the MDL for concentrations in sediment and/or biota can result in spurious and non-predictive BSAFs. In each case (chemical present below the MDL and chemical not detected at all), the resulting values must be flagged and different flags should be used for each case.

When plotting of the different C_{soc}-C_{ℓ} pairs is done, different symbols/colors should be used for the above two flagged data types. Examine this plot to see if the flagged data aligns with the general trend of the C_{soc}-C_{ℓ} pairs that are not flagged. Chemicals not

detected at all and chemicals with concentrations below the MDL should each be treated separately. One probably has greater confidence in the uncensored flagged data (below the MDL) than the chemicals not detected at all. This comparison/evaluation should be performed by doing the regression analysis without the flagged data, with the less-than-the-MDL flagged data included, and with the flagged data alone. Significance testing of the slopes (asking whether the slopes are different) should be done and these comparisons should help in determining whether to include or exclude the flagged data in the final regression. Examination of the residual plots should be done and will help greatly in determining whether to include or exclude chemicals present at concentrations less than MDL and/or chemicals not detected at all.

In general, for chemicals not detected at all (i.e., ½ of the MDL is used), they should be excluded from the analysis since these values are highly uncertain relative to the other C_{soc}-C_{ℓ} pairs. Additionally, the flagged data would, in high likelihood, be from sampling locations where less contamination existed and not the site of planned active remediation.

The above discussion was centered on non-detects and their use in the regression analysis. There are statistical approaches for averaging with censored data, i.e., non-detects (El-Shaarawi and Dolan, 1989; Newman et al., 1989; Newman, 1995). These approaches can be used with normally and log-normally distributed data. It is recommended that unbiased means be calculated only if less than 20% of the reported values are reported as being non-detect (Berthouex and Brown, 1994).

How do I estimate the confidence interval around a prediction?

The standard error of the geometric mean regression slope can be approximated by the standard error of the linear least-squares regression slope (Sokal and Rohlf, 1995). Most linear least-squares regression programs (SAS) or spreadsheets (Louts123 and Excel) calculate the standard error of the slope.

The 95% confidence limits on the slope would be calculated using student-t value:

$$Upper\ 95\%\ CI = b + s_b \times t_{0.05[n-2]} \quad \textbf{(13)}$$

$$Lower\ 95\%\ CI = b - s_b \times t_{0.05[n-2]} \quad \textbf{(14)}$$

where b is the geometric mean regression slope, s_b is the standard error of the geometric mean regression slope, n is the total number of data points used in the geometric mean regression, and $t_{0.05}$ is the two tailed Student-t for an $\alpha = 0.05\%$.

When calculating the geometric mean of the ratios of the C_{soc}-C_ℓ pairs (i.e., BSAFs), the averaging process in log space provides the mean and standard deviation. The 95% confidence limits would be calculated in log space using the mean and standard deviation, and then, the CIs would be transformed into arithmetic space. In arithmetic space, the 95% CI will be asymmetric.

Do I normalize by organic carbon and lipid?

Yes. The BSAF is the ratio of the concentration in the biota on a lipid basis to the concentration in the sediment on an organic carbon basis.

By working with C_{soc}-C_{ℓ} pairs (which are organic carbon and lipid normalized), one places these concentrations on a thermodynamic basis. By expressing the concentrations on a thermodynamic basis, the concentrations of the chemicals in sediment and tissue are corrected for differences in bioavailability and partitioning behavior. By using the thermodynamic based expressions, the C_{soc}-C_{ℓ} pairs are expressed equivalently.

Do I use weighted regression?

There are two general cases. First, when the C_{soc} and C_{ℓ} are individual observations (not averages), then individual C_{soc}-C_{ℓ} pairs should be given equal weights. Second, if the C_{soc} and C_{ℓ} are averages, then individual C_{soc}-C_{ℓ} pairs should be given equal weights except if the C_{soc} and C_{ℓ} variances are highly heterogeneous ($p<0.001$). If the variances are highly heterogeneous (very dissimilar), then perform both weighted (by the inverse of the variance) and unweighted regression and compare slopes. The heterogeneous variances might or might not have any appreciable effect on the slope. If appreciable effects exist on the slope, then the weighted regression model is preferred.

If I transform the data, do I need to use weighted regression?

See answer to previous question. The variances would need to be evaluated in log space for heterogeneity.

How do I take into account the home range of the biota whose tissue I measured?

As explained in the background, one must have knowledge of the organism's home range. With this information, sediment samples across the home range must be collected and analyzed, and the sediment samples must be representative of the organism's immediate life history. Accounting for the home range of the organism is done by

averaging the analytical results for sediment samples collected within the organism's home range.

What if my r^2 is low and my data do not plot with the appearance of an increasing linear function?

When this type of behavior is observed in the plot of C_{soc}-C_{ℓ} pairs, this is an extremely strong suggestion that different sampling locations have the different underlying conditions and parameters; e.g., different food webs, different organism populations, differences in chemical bioavailability, different diets, etc.; or a very limited dynamic range. In these cases, one will need to determine the factors causing these differences. If one can not resolve these difference, the same problems will also exist with other methods for predicting chemical residues, e.g., food web models, because these methods require this knowledge as well. In general, when this type of behavior is observed, the problem is in the data itself, and no statistical analysis method will circumvent the problem. Without resolving these differences, their effects will be reflected or incorporated into all calculations with the data.

How do I deal with outliers?

There are a number of different types of outliers. First, if the chemical was not detected at all, and ½ of the MDL was used, one could easily set these values aside without much criticism, in essence, making the argument that one has low confidence in the values. Second, if the chemical was flagged as being below the MDL and the uncensored value is reported, treating these values as outliers and setting them aside would be much harder. You would have to determine what level of confidence you place on values below the

MDL. In general, uncensored data below the MDL is included in the analysis unless there is an overwhelming reason to excluded the data, e.g., some type of methodological bias in the analytical technique. Third, the C_{soc}-C_{ℓ} pair is very different from the general population of C_{soc}-C_{ℓ} pairs. In this situation, always make sure the data are not miscalculated, transposed, or misidentified, and ensure that no other type of methodological error is associated with the data. If the data pair appears to be correct, statistical techniques are available for the testing of outliers.

Snedocor and Cochran (1980, p 167-168) present a statistical method for linear regression where the regression is performed without the outlier, and then the outlier is tested as to whether it is within sampling error of the population. The test criterion is a t-value. Because the outlier is not chosen randomly, to ensure a 1- α confidence, the calculated t-value is compared to the t-value from the t-table using α'; where α' equals α divided by n. Probably values for testing for outliers should be generally conservative, e.g., α = 5% or α = 1%. With an n of 20, the critical t-value for an α of 5% would be found using an α' of 0.25% with the t-table.

SAS software, software for statistical analysis, provides outlier detection and testing algorithms within its regression model program.

Do I develop a separate regression for each compound in a mixture?

Yes. This is most desirable because individual chemicals have different chemical properties. The differing behavior is most often observed with PCBs where fish appear

to be slightly enriched with the higher chlorinated PCB congeners relative to the distribution existing in the sediments.

When the value of x (i.e., exposure point concentration in sediment) is uncertain (e.g.,when biota migrate), how do I account for this in my regression?

The best method of accounting for organism migration is to design your sampling plan for the organism such that the organisms are collected just before they migrate back out of the site. This approach maximizes time the organism spends at the site of interest, and provides the best estimate of the residue in the organism based upon the organism's exposure in its immediate home range at the site.

Sampling design simulations (Burkhard, 2003) for the measurement of BSAFs (or C_{soc}-C_{ℓ} paired observations for determinations of BSAFs) suggest that spatial variability in the concentrations of the chemical does not add large uncertainties into the measured BSAF beyond those caused by temporal variability of the chemical concentrations in the water. Further, random walk migration simulations suggested that BSAFs (or C_{soc}-C_{ℓ} paired observations for determinations of BSAFs) can be measured with low uncertainty even when extreme spatial concentrations exist at the field site, provided the measurements are performed in more contaminated locations of the site for higher K_{ow} chemicals, i.e., $>10^5$ (Burkhard, 2003). The requirement of performing the field measurements at the more contaminated locations within the site will limit the regression approach because the range of C_{soc}-C_{ℓ} pairs will be small (see second paragraph of the Recommendations).

If the organisms spend a very short time at the site, e.g., the fish migrate through the site in a few days to a week, determination of BSAFs is not recommended even though the BSAF can be measured. The sediments from the site would not be reflective of the fish's recent exposure history.

Are there ways to improve my study design knowing what I know now about regression?

First, the importance of collecting sediment samples that are reflective of the organism's immediate home range can not be overstated. Spending time and resources to better define the relationship of the organisms to the sediments will greatly decrease the uncertainty associated with the resulting BSAFs. In addition, predictions using food web models, both steady-state or dynamic, will greatly improve because of the improved knowledge on the underlying relationship between the sediment and organism.

Second, it is important that composite samples reflective of the biota at the site of interest be collected. Clearly, collection and analysis of more organisms will provide a better measure of the average residue in the biota. However, biota samples consisting of mixed age classes is not recommended, e.g., juvenile and adult minnows, or one-year-old and three-year-old largemouth bass. Minimizing the differences in age (or size) will improve the quality of the biota samples and ultimately provide smaller variances for the biota residues. Typically, fishes of given size (e.g., smallest fish $\geq$75% of the largest fish) or age group (e.g., 3-year-olds) are collected.

After sample collection and analysis, plans should be made to visually examine the data by making plots of C_{soc}-C_{ℓ} paired observations and plots of BSAFs against C_{soc}. The C_{soc}s, C_{ℓ}s, and BSAFs should be plotted on a GIS type plot to determine if the values are correlated with geographical trends and conditions, e.g., the BSAFs increase with increasing distance away from the source on a river. Any additional information or understanding one can glean for the site will be advantageous in the remediation decision process.

As part of the overall study plan for successfully measuring a BSAF, time and resources should be allocated for resolving causes of non-linearity (when they exist) in the C_{soc}-C_{ℓ} paired observations. Resolving why will greatly aid in understanding the complexities of the site, and provide decision makers and risk assessors a much better basis for assessing and evaluating remediation options.

Deriving a BSAF using regression analysis or by calculating the average of the individual BSAFs uses the same data. Hence, it is suggested that BSAFs be derived using both approaches. The added effort for the second analysis should be relatively small since much of the effort, in performing the data analyses, is organizing the data into a usable form for the calculations.

REFERENCES

There are many standard college level textbooks on statistical analysis which include regression analysis. Almost all include discussion and examples on the linear least-squares

regression technique. Coverage of geometric mean regression analysis technique is often not addressed in standard college level textbooks. Halfon (1985) is an excellent reference on geometric mean regression. Sokal and Rohlf (1995) address the subject of model II regression including geometric mean regression.

Ankley, G.T., P.M. Cook, A.R. Carlson et al. 1992. Bioaccumulation of PCBs from sediments by oligochaetes and fishes: Comparison of laboratory and field studies. Can. J. Fish. Aquat. Sci. 49:2080-2085.

Berthouex, P.M. and L.C. Brown. 1994. Statistics for Environmental Engineers. Lewis Publishers/CRC Press, Boca Ration, FL.

Bierman, V.J., Jr. 1990. Equilibrium partitioning and biomagnification of organic chemicals in benthic animals. Environ. Sci. Technol. 24:1407-1412.

Burkhard, L.P. 2003. Factors influencing the design of bioaccumulative factor and biota-sediment accumulation factor field studies. Environ. Toxicol. Chem. 22(2):351-360.

Burkhard, L.P., P.M. Cook and D.R. Mount. 2003. The relationship of bioaccumulative chemicals in water and sediment to residues in fish: A visualization approach. Environ. Toxicol. Chem. 22(11):2822-2830.

Burkhard, L.P., P.M. Cook and M.T. Lukasewycz. 2004. Biota-sediment accumulation factors for polychlorinated biphenyls, dibenzo-p-dioxins, and dibenzofurans in southern Lake Michigan lake trout (*Salvelinus namaycush*). Environ. Sci. Technol. 38(20):5297-5305.

DiToro, D.M., C.S. Zarba, D.J. Hansen et al. 1991. Technical basis for establishing sediment quality criteria for nonionic organic chemicals using equilibrium partitioning. Environ. Toxicol. Chem. 10:1541-1583.

El-Shaarawi, A.H. and D.M. Dolan. 1989. Maximum likelihood estimation of water quality concentrations from censored data. Can. J Fish Aquat. Sci. 46(6):1033-1039

Ferraro, S.P., H. Lee Jr., R.J. Ozretich and D.T. Specht. 1990. Predicting bioaccumulation potential: a test of a fugacity-based model. Arch. Environ. Contam. Toxicol. 19(3):386-394.

Halfon, E. 1985. Regression method in ecotoxicology: A better formulation using the geometric mean functional regression. Environ. Sci. Technol. 19:747-749

Lake, J.L., N. Rubinstein and S. Pavignano. 1984. Predicting bioaccumulation: Development of a simple partitioning model for use as a screening tool in regulating ocean disposal of wastes. In: Fate and Effects of Sediment-Bound Chemicals in Aquatic Systems, K.L. Dickson, A.W. Maki and W.A. Brungs, Ed. Pergamon Press, New York, NY. p. 151-166.

Lake, J.L., N. Rubinstein, H. Lee II, C.A. Lake, J. Heltshe and S. Pavignano. 1990. Equilibrium partitioning and bioaccumulation of sediment-associated contaminants by infaunal organisms. Environ. Toxicol. Chem. 9:1095-1106.

McElroy, A.E. and J.C. Means. 1988. Factors affecting the bioavailability of hexachlorobiphenyls to benthic organisms. In: Aquatic Toxicology and Hazard Assessment, Vol. 10, W.J. Adams, G.A. Chapman and W.G. Landis, Ed. American Society for Testing and Materials, Philadelphia, PA. p. 149-158.

McFarland, V.A. and J.U. Clarke. 1986. Testing bioavailability of polychlorinated biphenyls from sediments using a two-level approach. In: Proceedings of the US Army Engineers Committee on Water Quality, 6th Seminar, R.G. Wiley, Ed. Hydraulic Engineering Research Center, Davis, CA. p. 220-229.

Minns, C.K. 1995. Allometry of home range size in lake and river fishes. Can. J. Fish. Aquat. Sci. 52:1499-1508.

Mood, A.M., F.A. Graybill and D.C. Boes. 1974. Introduction to the Theory of Statistics, 3rd ed. McGraw-Hill, New York, NY.

Newman, M.C. 1995. Quantitative Methods in Aquatic Ecotoxicology. Lewis/CRC Press, Boca Raton, FL.

Newman, M.C., P.M. Dixon, B.B Looney and J.E. Pinder III.. 1989. Estimating mean and variance for environmental samples with below detection limit observations. Water Res. Bull. 25(4):905-916.

Ricker, W.E. 1973. Linear regression in fishery research. J. Fish. Res. Board Can. 30:409-434.

Snedecor, G.W. and W.G. Cochran. 1980. Statistical Methods, 7th ed. Iowa State University Press, Ames, IA. p. 167-168.

Sokal, R.R. and F.J. Rohlf. 1995. Biometry: The Principles and Practice of Statistics in Biological Research, 3rd ed. W.H. Freeman and Co., New York, NY.

Thomann, R.V., J.P. Connolly and T.F. Parkerton. 1992. An equilibrium model of organic chemical accumulation in aquatic food webs with sediment interaction. Environ. Toxicol. Chem. 11:615-629.

Tracy, G.A. and D.J. Hansen. 1996. Use of biota-sediment accumulation factors to assess similarity of nonionic organic chemical exposure to benthically-coupled organisms of differing trophic mode. Arch. Environ. Contam. Toxicol. 30(4):476- 475.

Wong, C,S., P.D. Capel and L.H. Nowell. 2001. National-scale, field based evaluation of the biota-sediment accumulation factor model. Environ. Sci. Technol. 35(9):1709-1715.

APPENDIX

ECOLOGICAL RISK ASSESSMENT SUPPORT CENTER REQUEST FORM

Problem Statement: What is the most appropriate method to estimate the Biota Sediment Accumulation Factor (BSAF) from paired observations of concentrations in biota and sediment?

Requestors: Sharon Thoms and Al Hanke, Region 4

Background: BSAF is a parameter describing bioaccumulation of sediment-associated organic compounds or metals into tissues of ecological receptors. In a typical experiment to measure bioaccumulation the researcher collects colocated sediments and tissues over a gradient of contamination. Simple compared to bioaccumulation and trophic transfer models, it finds its use at Superfund sites to estimate progress toward achieving a protective tissue concentration as sediments become cleaner.

Expected Outcome: The expected outcome is a white paper addressing the following questions regarding the use of regression to obtain the most accurate estimate of BSAF:

Do I fit a straight line?
Do I plot my data on a log-log scale?
Do I force the line through the origin?
How do I handle non-detects?
How do I estimate the confidence interval around a prediction?
Do I normalize by organic carbon and lipid?
Do I use weighted regression?
If I transform the data, do I need to use weighted regression?
How do I take into account the home range of the biota whose tissue I measured?
What if my r^2 is low and my data do not plot with the appearance of an increasing linear function?
How do I deal with outliers?
Do I develop a separate regression for each compound in a mixture?
When the value of x (i.e., exposure point concentration in sediment) is uncertain (e.g., when biota migrate), how do I account for this in my regression?
Are there ways to improve my study design knowing what I know now about regression?

Where the topics are covered by standard books or web sites on statistics, they may be referenced. A few case studies may be useful to illustrate the concepts.

Additional Comments: Requestor can provide case studies.