



Analysis of Uncertainty in TBP Estimation of PAH Bioaccumulation Potential in Sediments

PURPOSE: Theoretical Bioaccumulation Potential (TBP) is commonly estimated to evaluate dredged sediments for open-water disposal. At present, such estimations are made without a measure of the uncertainty associated with the result. The purpose of this technical note is to demonstrate methods of calculating uncertainty in TBP.

BACKGROUND: Dredged sediment evaluations conducted under Corps of Engineers and EPA regulatory programs can require assessment of potential contaminant bioaccumulation from dredged material. Before laboratory bioaccumulation testing is deemed necessary, an equilibrium partitioning-based screening test called Theoretical Bioaccumulation Potential (TBP) is usually performed. The TBP model estimates the probable concentration of a neutral organic chemical that would eventually accumulate in an organism from continuous exposure to a sediment. TBP is calculated from chemical concentration and organic carbon content of the sediment, lipid content of target organisms, and relative affinities of the chemical for sediment organic carbon and animal lipid. Use of empirical biota-sediment accumulation factors for the relative affinity term has improved TBP estimation in recent years. However, routine calculation of TBP has not included analysis of uncertainty. This technical note demonstrates two methods for uncertainty analysis of TBP: an analytical method that incorporates random and systematic error, and a numerical method using bootstrap resampling of the model input parameters to calculate statistical uncertainty measures.

INTRODUCTION: The assumptions of the TBP model derive from thermodynamics: the system consisting of sediment, organism, and water is visualized as being closed. A neutral organic chemical in the system is given free movement and will distribute throughout the phases in contact until a condition of equilibrium is established. The concentrations at equilibrium are determined by chemical potential in each phase. Organic carbon in the sediment and lipid in the organism are assumed to be the primary compartments that account for partitioning of neutral chemicals. Thus, the concentration of a chemical in the sediment is normalized on the basis of its organic carbon content, and the application of a partition coefficient enables calculation of the expected equilibrium concentration in an exposed organism of stated lipid content (McFarland 1984, McFarland and Clarke 1986). The model equation is:

$$\text{TBP} = \text{BSAF} (C_s / \text{fOC}) \text{fL} \quad (1)$$

where the partition coefficient is the biota/sediment accumulation factor (BSAF), C_s is the concentration of neutral organic chemical in sediment, fOC is the decimal fraction total organic carbon content of the sediment, and fL is the decimal fraction lipid content of the target organism. Empirically derived BSAFs are calculated as:

$$\text{BSAF} = (C_t / \text{fL}) / (C_s / \text{fOC}) \quad (2)$$

DTIC QUALITY INSPECTED 4

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

19990902 115

where C_t / fL is the lipid-normalized contaminant concentration in the tissues of the exposed organism and C_s / fOC is the organic carbon-normalized contaminant concentration in the sediment to which the organism has been exposed.

TBP is a simplification of a complex and dynamic system that attempts to make predictions based on minimal sets of observations. TBP estimations, therefore, involve an undefined degree of uncertainty, incorporating both random and systematic error. Systematic error arises from inaccuracies or simplifications in the model equation, such as the effect on tissue residues of metabolism, growth of the organism, reproductive state, feeding behavior, and numerous other variables not included in the model. Random error is due to inaccuracies in the measurement of model input parameters arising from sources that are unknown and uncontrollable. Analysis of uncertainty can help to gauge the robustness of the model and the quality of the measurements that have been made, both in derivation of the BSAF used in the estimation and of the other TBP input parameters.

Numerical methods based on simulations are frequently used in uncertainty analysis for risk assessments. One such method is the bootstrap, which is particularly amenable to uncertainty analysis for simple models like TBP when replicate measurements are available for at least some of the model input parameters. The bootstrap belongs to a class of nonparametric, computer-intensive resampling techniques that are adaptable to a wide variety of statistical applications (Efron and Tibshirani 1986, 1991; Alden 1992; Alden and Rule 1992; Léger, Politis, and Romano 1992; Lutz et al. 1995; Basu, Guerra, and Read 1996; Jagoe and Newman 1997). The basic premise of the bootstrap is that the n replicate observations of some parameter X are the best indication of the unknown population distribution of X (Manly 1991). In the simplest application of the bootstrap, the n original observations of X can be resampled randomly with replacement to generate a large number of bootstrap samples, analogous to repeatedly shuffling a deck of cards and dealing hands. Each bootstrap sample can then be used to calculate a statistic such as the mean, and the standard error of that statistic is simply the standard deviation of the collection of bootstrap sample statistics. For a model such as TBP, the replicate observations for each input parameter in the TBP model are resampled many times, and the means from each set of bootstrap samples are then inserted in Equation 1 to generate numerous bootstrap estimates of TBP. From these, statistical uncertainty measures such as standard errors or confidence intervals can be calculated, or tests of significance can be conducted. Bootstrap tests of significance are free from distributional assumptions that are unlikely to hold for models consisting of derived ratio variables.

METHODS:

Experimental: BSAFs were determined (Equation 2) using pooled, field-exposed, lipid-normalized PAH tissue concentration data of seven benthic invertebrate taxa from the New York Bight Mud Dump Reference Site, and pooled organic-carbon normalized PAH concentration data of four composited sediment replicates from the same site (McFarland, Lutz, and Reilly 1994). In a separate study, sediments from Central San Francisco Bay, CA (Reference), and from Oakland Inner Harbor, CA (Oakland), were analyzed for PAHs and total organic carbon (TOC). Benthic clams (*Macoma nasuta*) and blue mussels (*Mytilus edulis*) from an uncontaminated area were analyzed for lipid. TBPs were calculated using Equation 1 to estimate potential for uptake of 15 PAHs and total PAH in the clams and mussels from exposure to the Oakland and Reference

sediments. Actual PAH bioaccumulation was measured following exposures of the bivalves to the two sediments for 28 days under bedded and suspended sediment conditions (McFarland et al. 1994).

Data Reporting: Much of the sediment data, and tissue bioaccumulation data at the New York Bight site, were reported as less than the method detection limit (MDL). Concentrations between 0.1 MDL and MDL were quantitated and flagged as "J" by the analyst, indicating that the compound was present but the measured concentration was less certain than that of concentration data > MDL. "J" values were used in the calculation of TBP and uncertainty measures in the same way as data above MDL. All "U" values (unquantitated, < MDL) were assumed to be < 0.1 MDL and were set equal to 0.1 MDL as a conservative estimate for inclusion in the calculations. When all replicates of a treatment were < MDL, the data were excluded from any calculations.

Analytical Uncertainty Method: Uncertainty for TBP was calculated by an analytical method adapted from Campbell (1982), which first estimates method error and propagated error separately and then combines them to calculate total error. Method error is the systematic error inherent in the model equation that exists even though the input parameters may be measured exactly. Propagated error is the random error associated with inaccuracies in measurement (or estimation) of the input parameters that is carried forward in the estimate made using the model.

Method error was estimated for TBP using field data of Sanders (1995). BSAFs were first calculated for 10 individual PAHs and total PAH from concentrations in oysters and sediments and the sediment TOC content recorded at 28 sites in Murrells Inlet, NC. Sites with sediment TOC < 0.2 percent were not included, as TOC normalization is considered inappropriate below this level (Di Toro et al. 1991). The lipid content used was the mean of all oysters analyzed from Murrells Inlet. TBP was then calculated for PAHs in oysters at 30 sites in North Inlet, NC, using the Murrells Inlet BSAFs and North Inlet C_s , fOC, and fL. The calculated TBPs were then compared against measured PAH C_t in oysters collected at the North Inlet sites. Sixty-five percent of the North Inlet PAH C_t were reported as < MDL, and these were excluded leaving $n = 113$ observations for generation of the method error data set. Average method error was calculated using the average of the absolute percentage differences between each measured North Inlet PAH C_t and the TBP estimation at the same site:

$$AME = \sum | [100 - (TBP*100)/C_t] | / n \quad (3)$$

TBP method error (ME) was then estimated for each PAH/exposure/organism combination using the average method error:

$$ME = TBP * AME/100 \quad (4)$$

Propagated error was determined separately for each of the four estimated input parameters of BSAF plus the three additional parameters in the TBP model, and then combined over all input parameters using Campbell's method 2 (Campbell 1982) for multiple input values. The following steps were involved:

- The mean (x) and lower and upper 95-percent confidence limits (CL_L and CL_U) of x were calculated for each input parameter.
- TBP at each limit (TBP_L and TBP_U) was calculated for each input parameter using CL_L and CL_U , respectively, for that parameter, and the means of all other input parameters.
- Propagated error was calculated for each input parameter using the formula

$$PE_i = ((TBP_L - TBP_U)^2 / 4)^{1/2} \quad (5)$$

- Overall propagated error was calculated:

$$PE = (\sum PE_i^2)^{1/2} \quad (6)$$

Total error in the estimation of TBP for each PAH/exposure/organism combination was calculated by evaluating method and propagated error simultaneously (Campbell 1982) as:

$$TE = (ME^2 + PE^2)^{1/2} \quad (7)$$

with total error considered an estimate of TBP uncertainty.

Numerical Uncertainty Method: For each PAH/exposure/organism combination, balanced bootstrap samples (Léger, Politis, and Romano 1992) were generated for TBP by combining 1,024 identical copies of the n replicate observations for each input parameter (i.e., doubling the data set 10 times) into a single large distribution, and then taking 1,024 random permutations of size n from the distribution and calculating the mean for each sample. The 1,024 sample means for each input parameter were substituted into Equation 1 to produce 1,024 estimates of TBP. A TBP bootstrap distribution was constructed by duplicating these estimates 6 times. The TBP bootstrap distribution was then randomly split into 1,024 TBP bootstrap samples of size $n = 6$. Mean TBP was calculated for each bootstrap sample, and the standard error of the mean (SE) was the standard deviation of the 1,024 bootstrap TBP means. The bootstrap coefficient of variation (CV) was equal to $SE(\sqrt{6}) / \text{mean}$. Lower and upper 95-percent confidence limits for TBP were defined as the 2.5th and 97.5th quantiles of the distribution of bootstrap means, using the simple bootstrap percentile method (Efron and Tibshirani 1986).

Measured tissue concentration data for each PAH/exposure/organism combination were also duplicated 1,024 times and randomly split into 1,024 balanced bootstrap bioaccumulation samples of size $n = 6$. A test of significance was performed by first calculating the difference $d_1 =$ the overall mean bootstrap TBP minus the original bioaccumulation mean for a given PAH/exposure/organism combination. The 1,024 bootstrap TBP samples and the 1,024 bootstrap bioaccumulation samples were then combined into a single distribution, shuffled, and randomly split into sets of 12 observations, the first 6 of which were assigned to TBP and the next 6 to measured bioaccumulation. The difference d was calculated as the mean bootstrap TBP minus the mean bootstrap bioaccumulation for each set of bootstrap sample comparisons. Finally, d_1 was added to the distribution of d from the 1,024 bootstrap sample comparisons for each PAH/exposure/organism combination, and all of the differences were ranked in ascending order. If d_1 fell within the upper 2.5 percent of the distribution of d (rank > 1,000), TBP was considered to significantly overestimate

the actual PAH tissue concentrations. If d_1 fell within the lower 2.5 percent of the distribution of d (rank < 26), TBP was considered to significantly underestimate the actual PAH tissue concentrations.

Sensitivity analysis was conducted for each TBP input parameter by calculating a minimum and maximum TBP using the minimum and maximum observed values of that parameter and the means of the other input parameters. A sensitivity index was computed as the absolute value of (max TBP - min TBP) / mean TBP for each parameter, with larger values of the index indicating greater contribution of the associated parameter to the uncertainty in TBP. The analysis was repeated for each of the eight treatments.

All computations were performed using SAS (SAS Institute, Inc. 1988).

RESULTS: PAH TBPs were calculated for the eight treatment combinations of clams and mussels exposed to bedded and suspended Oakland and Reference sediments. Results for clams exposed to suspended Oakland sediment are used in this paper to illustrate the uncertainty methods, and results for all treatments are summarized where appropriate. BSAFs from the New York Bight data, along with C_s , C_t , and calculated TBPs for the 15 PAHs and total PAH in Oakland sediment are given in Table 1.

Table 1						
TBP Input Parameters and Calculated TBPs for PAH Bioaccumulation in Clams Exposed to 50 mg/L Suspended Oakland Sediment (Concentrations in ng/g)						
PAH	BSAF	C_s mean (SE)*		C_t mean (SE)**		TBP†
Acp ^{††}	0.0158	1240	(111)	39.6	(10.5)	54.2
Acy	0.00764	69.3	(10.8)	0.37	(0.174)	1.47
Ant	0.0198	1770	(382)	215	(39.5)	97.0
B[a]A	0.181	2410	(259)	491	(39.8)	1210
B[a]p	0.0630	4310	(338)	379	(33.5)	752
B[b+k]F	0.0922	7370	(602)	718	(64.1)	1880
B[ghi]p	0.0589	3260	(306)	117	(12.5)	532
Chr	0.216	3200	(327)	697	(54.2)	1920
DBA	0.0210	432	(40.2)	26.6	(1.84)	25.2
Fla	0.222	7120	(661)	1870	(181)	4380
Flu	0.0147	534	(79.9)	25.9	(5.72)	21.7
l[cd]P	0.0464	3600	(300)	85.2	(9.83)	463
Naph	0.118	550	(38.0)	53.6	(6.13)	180
Phen	0.278	5050	(459)	610	(111)	3890
Pyr	0.360	7330	(651)	1960	(194)	7300
Tot	0.0786	48200	(3640)	7290	(659)	10500

* dry wt., $n = 5$
 ** wet wt., $n = 6$
 † fOC mean (SE) = 0.0111 (0.00085) dry wt., $n = 5$; fL mean (SE) = 0.030748 (0.00493) wet wt., $n = 6$
 †† Acenaphthene, acenaphthylene, anthracene, benz[a]anthracene, benzo[a]pyrene, benzo[b+k]fluoranthene, benzo[ghi]perylene, chrysene, dibenzanthracene, fluoranthene, fluorene, indeno[123cd]pyrene, naphthalene, phenanthrene, pyrene, total of 15 PAH

Analytical Analysis of Uncertainty: TBP model validation using the data of Sanders (1995) in Equation 3 resulted in average model uncertainty of 44 percent of the measured value. The maximum model uncertainty was 313 percent and the minimum was 5 percent.

Estimated uncertainty expressed as propagated, method, and total error for the calculated PAH TBPs in clams exposed to suspended Oakland sediment is summarized in Table 2, and an example of the propagated error calculations for TBP is given in Table 3. The largest propagated errors were associated with the New York Bight C_t and fOC components of BSAF. In the example, overall propagated error was 159 percent of the calculated TBP for acenaphthene in clams exposed to suspended Oakland sediment, and total error was 165 percent of TBP. Propagated error was always at least twice the method error. Total error always exceeded the value of TBP and averaged 1.8 times TBP. For all eight treatment combinations included in this experiment, PAH total error ranged from 1.3 to 4.8 times the value of TBP. Error percentages approaching or exceeding the value of TBP are not unexpected, given the compounded errors associated with the multiple input parameters comprising TBP.

Table 2
Uncertainty Expressed as Propagated, Method, and Total Error for Calculated TBP of PAHs in Clams Exposed to 50 mg/L Suspended Oakland Sediment (Concentrations in ng/g*; abbreviations of PAHs from Table 1)

PAH	Propagated Error	Method Error	Total Error	Total Error as Percent of TBP	TBP/ C_t
Acp	86.4	23.8	89.6	165	1.4
Acy	2.03	0.65	2.13	145	4.0
Ant	165	42.7	170	176	0.4
B[a]A	1830	531	1910	158	2.5
B[a]p	1000	331	1060	141	2.0
B[b+k]F	2550	828	2680	142	2.6
B[ghi]p	713	234	750	141	4.5
Chr	2880	845	3000	156	2.8
DBA	35.3	11.1	37.0	147	1.0
Fla	21100	1930	21200	483	2.3
Flu	35.8	9.54	37.1	171	0.9
I[cd]P	648	204	679	147	5.4
Naph	215	79.2	229	127	3.4
Phen	6506	1710	6730	173	6.4
Pyr	21400	3210	21600	296	3.7
Tot	14000	4620	14800	140	1.4

* Wet weight, $n = 6$

Numerical Analysis of Uncertainty: From the 1,024 bootstrap estimates of TBP for each PAH, the bootstrap mean TBP, SE, CV, and 95-percent CI were calculated; these are presented in Table 4 for clams exposed to suspended Oakland sediment. The bootstrap mean TBPs were positively biased compared with the calculated TBPs (Table 1). The amount of bias ranged from 2-3 percent of the calculated TBP for most PAHs to 20 percent of the calculated TBP for pyrene, and >60 percent of the calculated TBPs for benz[a]anthracene and chrysene. The positive bias is reflective of skewness

Table 3 Example of Propagated Error Calculations: Acenaphthene TBP for Clams Exposed to 50 mg/L Suspended Oakland Sediment (Concentrations in ng/g (C_t and C_s) or Decimal Fraction (fL and fOC))						
Input	Mean (\bar{x})	CL_L	CL_U	TBP_L	TBP_U	PE_i
BSAF:						
NY C_t	3.56	-0.171	7.28	-2.60	111	56.7
NY fL	0.0267	0.0197	0.0337	73.4	42.9	15.2
NY C_s	42.5	34.5	50.5	66.6	45.6	10.5
NY fOC	0.0050	-0.0001	0.0102	-1.27	110	55.4
Oakland C_s	1240	931	1550	40.7	67.6	13.5
Oakland fOC	0.0111	0.0087	0.0135	68.7	44.7	12.0
Oakland fL	0.0307	0.0181	0.0434	31.8	76.5	22.3
Overall propagated error						86.4

Table 4 Statistical (bootstrap) Error for TBP Estimated for PAHs in Clams Exposed to 50 mg/L Suspended Oakland Sediment (Concentrations in ng/g; abbreviations of PAHs are from Table 1)					
PAH	Bootstrap Mean TBP	Bootstrap SE	Bootstrap CV	Bootstrap 95 Percent CI	CI Width as Percent of Bootstrap Mean TBP
Acp	55.8	15.3	0.61	(31.2, 91.6)	108
Acy	1.50	0.31	0.47	(0.99, 2.22)	82
Ant	99.0	27.3	0.62	(57.3, 163)	107
B[a]A	1960	856	0.98	(828, 4040)	164
B[a]p	777	160	0.46	(496, 1100)	78
B[b+k]F	1930	383	0.44	(1250, 2750)	78
B[ghi]p	543	111	0.46	(361, 789)	79
Chr	3160	1490	1.05	(1380, 6740)	169
DBA	25.8	5.69	0.49	(16.7, 39.6)	89
Fla	4940	1460	0.66	(2760, 8300)	112
Flu	22.1	5.98	0.60	(12.3, 35.8)	106
I[cd]P	473	101	0.48	(300, 692)	83
Naph	185	29.2	0.35	(135, 249)	62
Phen	4130	1020	0.55	(2460, 6410)	96
Pyr	8740	2700	0.69	(4790, 14900)	116
Tot	10700	2070	0.43	(7060, 15200)	76

to the right in the bootstrap distribution of TBP. For cases of extreme skewness, calculating the bootstrap TBP using medians reduces the bias, e.g. to -2 percent for pyrene and 6 percent for benz[a]anthracene and chrysene. Bootstrap CVs ranged from 0.31 to 1.07 for the eight treatment combinations, and bootstrap 95 percent CI width as a measure of uncertainty was a factor of 0.5 to 1.9 times the bootstrap mean TBP.

Bootstrap Tests of Significance: The bootstrap mean TBPs and their 95-percent confidence intervals are shown with the measured PAH concentrations in clam tissues of the suspended Oakland sediment treatment in Figure 1, and Table 5 presents the bootstrap test of significance for that treatment. In this example, TBP significantly overestimated bioaccumulation of eight individual PAHs and significantly underestimated bioaccumulation of only one PAH. Over all eight treatment combinations, TBP significantly overestimated PAH bioaccumulation in 41 percent of cases, significantly underestimated PAH bioaccumulation in 10 percent of cases, and did not differ significantly from PAH bioaccumulation in 49 percent of cases. Conducting a bootstrap test of significance on the difference in median bootstrap PAH TBPs and median PAH tissue concentrations resulted in only 4 percent significant overestimates and 6 percent significant underestimates among the 16 PAHs and 8 treatment combinations, thus providing a more conservative test than the bootstrap test of significance using means.

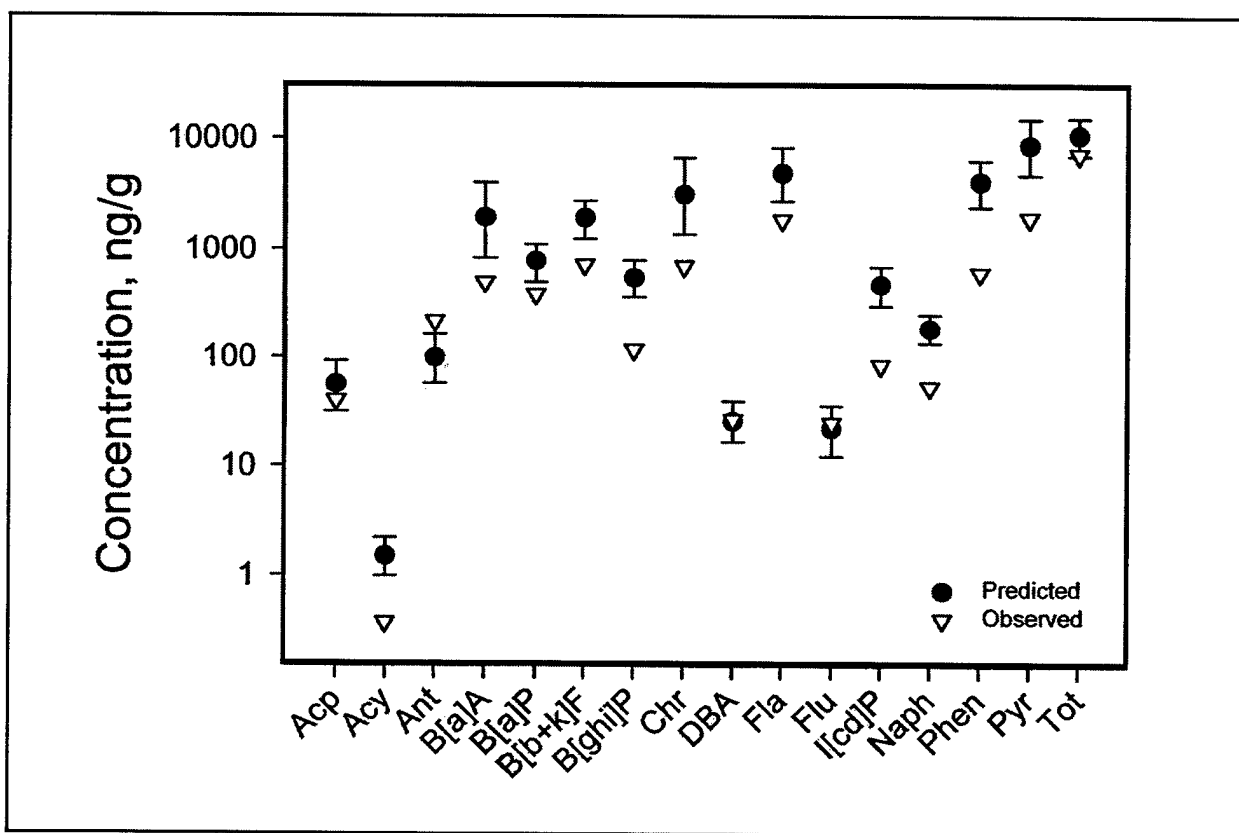


Figure 1. Measured concentrations of PAHs in tissues of clams, *Macoma nasuta*, exposed to suspended Oakland sediment, and bootstrap TBP-predicted concentrations. Bars are 95-percent confidence limits. Abbreviations of PAHs from Table 1

Figure 2 shows the overall correspondence between the predicted and observed tissue concentrations in the Oakland and Reference sediment treatments. TBP was more likely to overestimate or underestimate actual PAH tissue concentrations in the Reference sediment treatments than in the more contaminated Oakland sediment treatments.

Table 5
Bootstrap Test of Significance for the Difference in Bootstrap Mean TBP and Mean Measured Tissue Concentration of PAHs in Clams Exposed to 50 mg/L Suspended Oakland Sediment (Concentrations in ng/g; abbreviations of PAHs from Table 1)

PAH	Bootstrap Mean TBP	Mean Measured Conc.	Difference in Means, d_1	Rank of d_1 in Bootstrap d Distribution	Bootstrap Probability P^*
Acp	55.8	39.6	16.2	838	0.3665
Acy	1.50	0.37	1.14	1013	0.0253 O
Ant	99.0	216	-117	15	0.0292 U
B[a]A	1960	492	1470	974	0.1014
B[a]p	777	379	398	1001	0.0487 O
B[b+k]F	1930	719	1220	1014	0.0234 O
B[ghi]p	543	117	426	1020	0.0117 O
Chr	3160	697	2470	971	0.1072
DBA	25.8	26.6	-0.78	459	0.8947
Fla	4940	1870	3070	989	0.0721
Flu	22.1	24.9	-2.77	378	0.7368
I[cd]P	473	85.3	388	1020	0.0117 O
Naph	185	53.6	131	1019	0.0137 O
Phen	4130	610	3520	1020	0.0117 O
Pyr	8740	1960	6780	1004	0.0429 O
Tot	10700	7290	3420	945	0.1579

* TBP significantly overestimates (**O**) or underestimates (**U**) actual bioaccumulation, two-tailed $\alpha = 0.025$.

Sensitivity Analysis: Among the TBP input parameters, the major contributors to uncertainty were the BSAF parameters, particularly the New York PAH tissue concentrations and lipid fraction, each of which accounted for about 30 percent of TBP parameter uncertainty based on the calculated sensitivity index (Table 6). Together, the four BSAF parameters accounted for >80 percent of total TBP parameter uncertainty.

DISCUSSION: Determining method error requires that a database be available in which estimates made using the model can be compared against measured values (taken as the true values) of a set of chemicals. Method error in estimating some chemical properties, e.g., density, index of refraction, or heat capacity, can be as low as 1-2 percent. Other chemical property estimation method errors range from 3-20 percent, and some of the estimation methods for equilibrium partitioning relationships commonly used in chemical fate modeling, e.g., K_{ow} , BCF, K_{oc} , and H, have method errors of at least an order of magnitude (Lyman, Reehl, and Rosenblatt 1982). Large databases that can be used to assess method error for TBP do not yet exist. However, for TBP prediction of PAH bioaccumulation in bivalves it was possible to estimate method error using a recently published study on PAH contamination of oysters in two North Carolina estuaries (Sanders 1995). The two estuaries differed in degree of PAH contamination, one receiving urban run-off and the other receiving run-off from heavily forested terrain. For the purposes of this paper, data from the urban estuary were used to calculate BSAFs and data from the rural estuary were used along with those

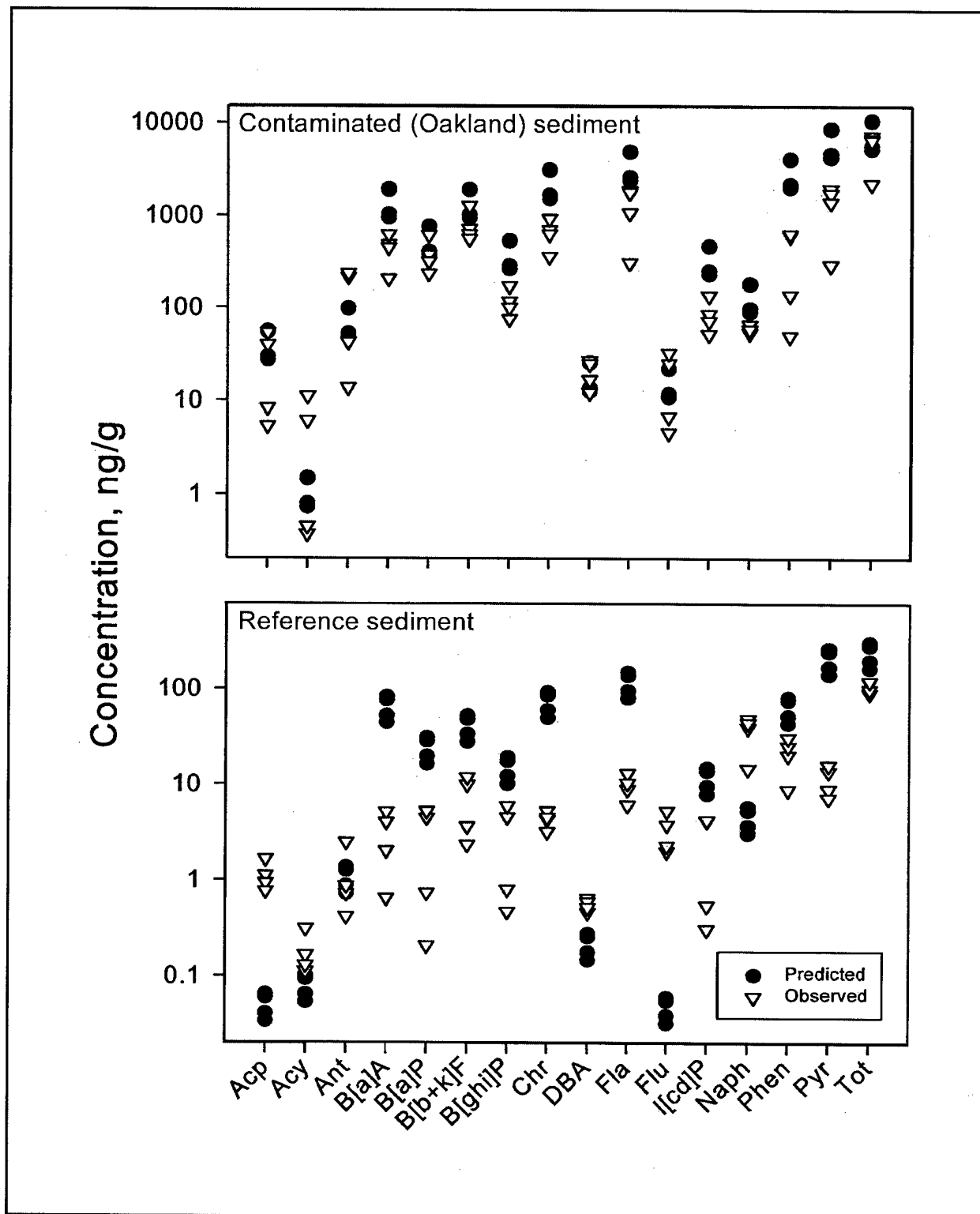


Figure 2. Measured concentrations and bootstrap TBP-predicted concentrations of PAHs in clams, *Macoma nasuta*, and mussels, *Mytilus edulis*, exposed to bedded and suspended sediments. Abbreviations of PAHs from Table 1

BSAFs to calculate TBPs. The TBP predictions were then compared with true (measured) tissue concentrations from the rural estuary to determine method error for TBP. The 44-percent average method error obtained is relatively low compared with most other estimated chemical fate modeling equilibrium-partitioning relationships. The reason for this is that for highly hydrophobic compounds, water-referenced partition coefficients such as K_{ow} require measurement of extremely low concentrations in the water phase, whereas BSAFs do not require knowledge of the water-phase concentrations.

Propagated error was, on average, fourfold greater than method error. In the example used in Table 3, and in general for the data not shown, the largest contributors to propagated error were the New York fOC and C_t components of the BSAFs. Sensitivity analysis based on maximum and minimum values of the TBP input parameters indicated that the largest contributors to parameter uncertainty were New York lipid and C_t . The magnitude of uncertainty in the New York BSAF parameters results from several factors: large variability in the C_t data, including high outliers for many of the PAHs; inclusion of estimated (< MDL) rather than quantitated concentration data for C_t and especially C_s ; and extremely low values (<1 percent) for the normalizing parameters TOC and lipid. Sediment TOC was only 0.5 percent and it has been observed that variability in the relationship between bioavailable and sediment-bound neutral chemicals increases rapidly as sediment TOC falls below 1.0 percent (McFarland et al. 1996). Problems arising from these various factors include the arbitrary decision of what value, if any, to substitute for the nonquantitated below-MDL observations for inclusion in the model calculations; the disproportionate influence on TBP of denominator input parameters with values close to zero; and the disproportionate effect of outliers in general on any derived variable.

The more complex a model is, and the greater the number of input values required that must be either estimated or measured, the greater the propagated error becomes. TBP is a relatively simple estimator, containing a total of seven input parameters (including the four that are required for BSAF). From the acenaphthene example of Table 3 and the sensitivity analysis (Table 6), it is clear that the quality of TBP estimation is highly dependent on the quality of the BSAF. For example, had the Table 3 New York C_t and fOC input data propagated errors been as low as those of the other input parameters (14-15 ng/g or decimal fraction, on average) the overall propagated error would have been about 40 ng/g rather than 86, and the total error would have been about 46 ng/g. The effect on the estimation would have been a decrease in total error from 165 percent of TBP to about 85 percent for the example given. Calculation of BSAFs using more precise data with fewer or no outliers or below-MDL values should improve the TBP estimator and reduce the magnitude of uncertainty.

Analysis of uncertainty for simple environmental models such as TBP, using either the analytical methodology or the numerical bootstrap methodology, permits an evaluation of the utility of these models for their intended purpose. For example, a model with associated uncertainty spanning orders of magnitude would have little value as a predictive or screening tool, and could be costly or counterproductive if applied in a regulatory context. Despite the problems inherent in the BSAF parameter data used in this study, TBP uncertainty was less than a factor of 2 for the majority of PAHs, and always less than a factor of 5 as estimated by the analytical method; and less than a factor of 2 in all cases as estimated by the bootstrap method.

Table 6
Sensitivity Analysis of TBP Parameters

TBP Input Parameter	Sensitivity Index*	
	Oakland Sediment	Reference Sediment
fOC	0.42	0.04
C _s	0.57	0.90
fL	0.34 - 1.29**	0.32 - 0.96**
BSAF:		
NY fOC	1.44	1.44
NY C _s	1.45	1.45
NY C _t	3.46	3.46
NY fL	3.72	3.72

* Average for 15 PAHs and total PAH.
 ** Range for four organism/treatment combinations.

Multi-parameter alternatives to the simple equilibrium partitioning model have been developed (Morrison et al. 1996, Burkhard 1998, Iannuzzi et al. 1996). One such model (Morrison et al.) uses fugacity calculations and considers diet, feeding behavior, chemical disequilibrium between sediment and water, and other factors for a total of about 40 input variables. The 95-percent confidence limits of estimates made using both a complex model and simple equilibrium partitioning with BSAFs were compared to demonstrate the superiority of the complex model. The 95-percent confidence interval was expressed as the factor that should be applied to the estimation in order to account for 95 percent of the observed data, and was calculated based on the log normal distribution of deviations between observed and predicted fugacity ratios (Table 4 in Morrison et al. (1996)). The range of factors for the complex fugacity model was 1.9 to 7, and if the high value of 7 was excluded, all factors were equal to 3.2 or less. Exclusion of the high value was considered appropriate because the model is most sensitive to diet-related parameters, and discrepancies between predicted and observed fugacity ratios (corresponding to BSAFs) may be due to unrealistic diet assignments. For equilibrium partitioning, the range calculated by Morrison et al. was higher, 2.5 to 9. A similar exercise using the differences in predicted and observed tissue concentrations from the eight treatment combinations of this study resulted in a range of 2.9 to 3.9 for the Reference sediment treatments and 3.1 to 5.6 for the Oakland sediment treatments.

Multi-parameter models developed by Gobas and Thomann use the octanol-water partition coefficient and other parameters to predict bioaccumulation factors (Burkhard 1998). Uncertainty for these models, based on the 10th to 90th percentile range, was reported as a factor of 3.3 to 5.5 for the Gobas model and 3.3 to 8.7 for the Thomann model. A probabilistic food web model developed by Iannuzzi et al. (1996) to predict bioaccumulation of xenobiotic chemicals from sediment produced distributional results that were generally within an order of magnitude or less of the range of concentrations of five coplanar PCB congeners measured in tidal Passaic River organisms. Uncertainty factors for this model were not reported.

TBP cannot be expected to serve as a precise estimator of bioaccumulation given the multitude of environmental factors that influence actual bioaccumulation of contaminants. However, TBP has been shown to have predictive ability similar to that of far more complex multi-parameter

bioaccumulation models. TBP should be sufficiently accurate to identify sediments that have negligible likelihood of causing unacceptable bioaccumulation of neutral organic chemicals, or, conversely, to signal when neutral organic contaminant bioaccumulation may be of environmental concern and laboratory bioaccumulation testing should be performed. TBP significantly underestimated actual PAH bioaccumulation only 10 percent of the time in this study, while it significantly overestimated PAH bioaccumulation 41 percent of the time. Thus, for screening purposes, TBP is a reasonably accurate to somewhat conservative predictor of bioaccumulation.

POINTS OF CONTACT: For additional information, contact one of the authors, Dr. Victor A. McFarland, (601) 634-3721, mcfarlv@wes.army.mil, or Ms. Joan U. Clarke, (601) 634-2954, clarkej@wes.army.mil, or the manager of the Long-Term Effects of Dredging Operations Program, Dr. Robert M. Engler, (601)634-3624, englerr@wes.army.mil. This technical note should be cited as follows:

McFarland, V. A., and Clarke, J. U. (1999). "Analysis of uncertainty in TBP estimation of PAH bioaccumulation potential in sediments," *Dredging Research Technical Notes Collection* (EEDP-04-32), U.S. Army Engineer Research and Development Center, Vicksburg, MS. www.wes.army.mil/el/dots/eedptn.html

REFERENCES

- Alden, R. W. III. (1992). "Uncertainty and sediment quality assessments: I. Confidence limits for the triad," *Environmental Toxicology and Chemistry* 11, 637-644.
- Alden, R. W. III, and Rule, J. H. (1992). "Uncertainty and sediment quality assessments: II. Effects of correlations between contaminants on the interpretation of apparent effects threshold data," *Environmental Toxicology and Chemistry* 11, 645-651.
- Basu, S., Guerra, R., and Read, R. (1996). "Bootstrap confidence intervals for concentration parameters in dilution assays," *Journal of Agricultural, Biological, and Environmental Statistics* 1, 454-466.
- Burkhard, L. P. (1998). "Comparison of two models for predicting bioaccumulation of hydrophobic organic chemicals in a Great Lakes food web," *Environmental Toxicology and Chemistry* 17, 383-393.
- Campbell, C. (1982). "Evaluating propagated and total error in chemical property estimates." *Handbook of Chemical Property Estimation Methods. Environmental Behavior of Organic Compounds*. W. J. Lyman, W. R. Reehl, and D. H. Rosenblatt, ed., McGraw-Hill Book Company, New York, Appendix C.
- Di Toro, D. M., Zarba, C. S., Hansen, D. J., Berry, W. J., Swartz, R. C., Cowan, C. E., Pavlou, S. P., Allen, H. E., Thomas, N. A., and Paquin, P. R. (1991). "Technical basis for establishing sediment quality criteria for nonionic organic chemicals by using equilibrium partitioning," *Environmental Toxicology and Chemistry* 10, 1541-1586.
- Efron, B., and Tibshirani, R. (1986). "Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy," *Statistical Science* 1, 54-77.
- Efron, B., and Tibshirani, R. (1991). "Statistical data analysis in the computer age," *Science* 253, 390-395.
- Iannuzzi, T. J., Harrington, N. W., Shear, N. M., Curry, C. L., Carlson-Lynch, H., Henning, M. H., Su, S. H., and Rabbe, D. E. (1996). "Distributions of key exposure factors controlling the uptake of xenobiotic chemicals in an estuarine food web," *Environmental Toxicology and Chemistry* 15, 1979-1992.
- Jago, R. H., and Newman, M. C. (1997). "Bootstrap estimation of community NOEC values," *Ecotoxicology* 6, 293-306.
- Léger, C., Politis, D. N., and Romano, J. P. (1992). "Bootstrap technology and applications," *Technometrics* 34, 378-398.

- Lutz, M. W., Kenakin, T. P., Corsi, M., Menius, J. A., Krishnamoorthy, C., Rimele, T., and Morgan, P. H. (1995). "Use of resampling techniques to estimate the variance of parameters in pharmacological assays when experimental protocols preclude independent replication: An example using Schild regressions," *Journal of Pharmacology and Toxicology Methods* 34, 37-46.
- Lyman, W. J., Reehl, W. F., and Rosenblatt, D. H. (1982). *Handbook of Chemical Property Estimation Methods. Environmental Behavior of Organic Compounds*, McGraw-Hill Book Company, New York.
- Manly, B. F. J. (1991). *Randomization and Monte Carlo Methods in Biology*. Chapman and Hall, New York.
- McFarland, V. A. (1984). "Activity-based evaluation of potential bioaccumulation from sediments," *Dredging '84 Proceedings*, Vol 1, Clearwater Beach, FL, Nov. 14-16, 1984, 461-467.
- McFarland, V. A. and Clarke, J. U. 1986. "Testing bioavailability of polychlorinated biphenyls from sediments using a two-level approach," *Proceedings, USACE Committee on Water Quality, Sixth Seminar*, New Orleans, LA, Feb. 25-27, 1986, 220-229.
- McFarland, V. A., Clarke, J. U., Lutz, C. H., Jarvis, A. S., Mulhearn, B., and Reilly, F. J., Jr. (1994). "Bioaccumulation potential of contaminants from bedded and suspended Oakland Harbor Deepening Project sediments to San Francisco Bay flatfish and bivalve mollusks." Miscellaneous Paper EL-94-7, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.
- McFarland, V. A., Honeycutt, M. E., Feldhaus, J., Ace, L. N., Brannon, J. M., Weiss, C. A., Clarke, J. U., McCant, D., and Jones, P. (1996). "Lower limits of organic carbon normalization: Results of fish/sediment/water equilibrium partitioning studies," Environmental Effects of Dredging Technical Notes EEDP-01-38, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.
- McFarland, V. A., Lutz, C. H., and Reilly, F. J., Jr. (1994). "Bioaccumulation data and analysis for selected contaminants in sediments and biota of the New York Bight Apex Mud Dump Reference Site." Final Report to U.S. Army Engineer District, New York. U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.
- Morrison, H. A., Gobas, F. A. P. C., Lazar, R., and Haffner, G. D. (1996). "Development and verification of a bioaccumulation model for organic contaminants in benthic invertebrates," *Environmental Science and Technology* 30, 3377-3384.
- Sanders, M. (1995). "Distribution of polycyclic aromatic hydrocarbons in oyster (*Crassostrea virginica*) and surface sediment from two estuaries in South Carolina," *Archives of Environmental Contamination and Toxicology* 28, 397-405.
- SAS Institute Inc. (1988). *SAS Procedures Guide, Release 6.03 Edition*. SAS Institute Inc., Cary, NC.

NOTE: The contents of this technical note are not to be used for advertising, publication, or promotional purposes. Citation of trade names does not constitute an official endorsement or approval of the use of such products.