U.S. Army
Corps of Engineers

New England District
Concord, Massachusetts

U.S. Environmental
Protection Agency

New England Region
Boston, Massachusetts

RESPONSIVENESS SUMMARY TO THE PEER REVIEW OF MODEL CALIBRATION:
MODELING STUDY OF PCB CONTAMINATION IN THE HOUSATONIC RIVER

DCN: GE-080105-ACUD

January 2006

Environmental Remediation Contract
GE/Housatonic River Project
Pittsfield, Massachusetts

Contract No. DACW33-00-D-0006

Task Order 0003

05P-0949-1

RESPONSIVENESS SUMMARY TO THE PEER REVIEW OF MODEL CALIBRATION: Modeling Study of PCB Contamination in the Housatonic River

ENVIRONMENTAL REMEDIATION CONTRACT
GENERAL ELECTRIC (GE)/HOUSATONIC RIVER PROJECT
PITTSFIELD, MASSACHUSETTS

Contract No. DACW33-00-D-0006
Task Order No. 0003

DCN: GE-080105-ACUD

Prepared for

U.S. ARMY CORPS OF ENGINEERS
New England District
Concord, Massachusetts

and

U.S. ENVIRONMENTAL PROTECTION AGENCY
New England Region
Boston, Massachusetts

Prepared by

WESTON SOLUTIONS, INC.
West Chester, Pennsylvania

January 2006

Work Order No. 20123.001.096.0680

AUTHORS/CONTRIBUTORS

Tony Donigian
AQUA TERRA Consultants
Mountain View, California

Gary Lawrence
EVS Environment Consultants, member of the Golder group of companies
Vancouver, British Columbia

Edward Garland, Paul Paquin
HydroQual, Inc.
Mahwah, New Jersey

Rich DiNitto, Richard McGrath
Sleeman Hanley & DiNitto, Inc.
Boston, Massachusetts

Susan Svirsky
U.S. Environmental Protection Agency
Boston, Massachusetts

Scott Campbell
Weston Solutions, Inc.
West Chester, Pennsylvania; Pittsfield, Massachusetts

Jonathan Clough
Warren Pinnacle Consulting, Inc.
Warren, Vermont

TABLE OF CONTENTS

Section **Page**

LIST OF ACRONYMS

ADCP	Acoustic Doppler Current Profiler
ANOVA	Analysis of Variance
BSAF	biota-to-sediment accumulation factor
cfs	cubic feet per second
cm/d	centimeters per day
CMS	Corrective Measures Study
DQO	data quality objective
EDS	energy-dispersive system
EPA	U.S. Environmental Protection Agency
EPC	exposure point concentration
EqP	equilibrium partitioning
FCM	Food Chain Model
GE	General Electric Company
IMPG	Interim Media Protection Goal
IQG	Information Quality Guideline
m	meter
m^2	square meter
µg/L	micrograms per liter
MDEP	Massachusetts Department of Environmental Protection
MFD	Modeling Framework Design
mg/kg	milligrams per kilogram
OC	organic carbon
PCB	polychlorinated biphenyl
POM	particulate organic matter
PSA	Primary Study Area
QA	quality assurance
QAPP	Quality Assurance Project Plan
RFI	Resource Conservation and Recovery Act Facility Investigation
SEM	scanning electron microscopy
SNL	Sandia National Laboratories
SOP	standard operating procedure
TEQ	toxic equivalence
TOC	total organic carbon
tPCBs	total polychlorinated biphenyls
TSS	total suspended solids

WOE	weight-of-evidence
WWTP	wastewater treatment plant
XRD	X-ray diffraction

INTRODUCTION

This document presents the response from the U.S. Environmental Protection Agency (EPA) to comments and questions raised by an independent Peer Review Panel following their review of the Model Calibration Report for the GE/Housatonic River Site Rest of River released in December 2004. The review was conducted by seven experts in the field of numerical modeling of aquatic and riverine systems. This document, referred to herein as the Model Calibration Responsiveness Summary, has been prepared as part of EPA's obligations under Paragraph 22.c and Appendix J of the comprehensive agreement relating to the cleanup of the General Electric Company (GE) Pittsfield, MA facility, certain off-site properties, and the Housatonic River (referred to as the "Consent Decree"). The Consent Decree was entered on October 27, 2000, by the United States District Court of Massachusetts - Western Division, located in Springfield, MA.

Under the terms of the Consent Decree, EPA was required to conduct modeling of the fate, transport, and bioaccumulation of polychlorinated biphenyls (PCBs) in the area referred to as the "Rest of the River," defined as the area of river and adjacent floodplain downstream from the confluence of the East and West Branches of the Housatonic River in Pittsfield, MA. The Consent Decree further stipulated that the models used will include a hydrodynamic component, a sediment transport component, a PCB fate and transport component, and a bioaccumulation component. Following completion of the Validation Report Peer Review, the model will be used by GE as a tool in comparing the relative effectiveness of proposed remedial alternatives, including baseline conditions.

Prior to the Peer Review, a public comment period provided the opportunity for the public and GE to submit written comments on the Model Calibration Report for consideration by the Peer Review Panel, within the context of the Peer Review Charge. On May 4 and 5, 2004, the Model Calibration Peer Review Panel ("Reviewers") met at a public forum in Lenox, MA, to review and discuss the Model Calibration Report within the framework of the Charge. During this meeting, the members of the public and GE were provided the opportunity to present oral comments to the Panel, and the Panel was able to engage in a question/answer session with the public presenters. The Reviewers subsequently submitted final written comments to EPA's Managing Contractor for the Peer Review, SRA International, Inc., of Arlington, VA. This document is EPA's formal response to the final written Peer Review comments.

APPROACH AND ORGANIZATION OF THIS DOCUMENT

As stipulated in Appendix J to the Consent Decree, Peer Reviewers were discouraged from discussing their individual comments with each other outside the public Peer Review Meeting, to allow the full discussion to take place in public. In addition, the Reviewers were not required to reach consensus; therefore, the comments were prepared independently by each Reviewer. As observed during the Peer Review itself, many of the Reviewers noted some of the same issues with the Model Calibration Report; therefore, they submitted similar written comments on these issues. Conversely, as might be expected, at many times Reviewers had differing views on issues; this is also reflected in the written comments.

As a result of these considerations, and to avoid unnecessary repetition and to increase clarity in the Model Calibration Responsiveness Summary, EPA organized this document so that responses to general issues are presented first, followed by responses to specific comments.

The first section, termed "Response to General Issues," addresses issues that were raised by a number of Reviewers and/or had broad implications for the model calibration and for the modeling study in general. In this first section, EPA has identified 12 General Issues and has provided a Summary of Issue statement for each to frame the technical basis for the issue and to provide an indication of how often the issue was noted by the Reviewers. Each Summary of Issue statement is followed by EPA's response to the General Issue. Many of the responses to specific comments from the Reviewers refer back to the responses to the General Issues.

The second section is entitled "Response to Specific Comments." In this section, each Reviewer's comments are repeated verbatim in their entirety, grouped according to the structure of the Model Calibration Report Peer Review Charge. Because some Reviewers also provided comments outside of the Charge questions, it was necessary to add a section entitled "Overview Comments" at the beginning of this section, and a section entitled "Additional Comments" at the end. In each subsection, the comments of the individual Reviewers are presented in alphabetical order, with responses from EPA inserted at appropriate intervals. As noted above, many of these comments refer the reader to one or more of the General Issues responses.

RELATIONSHIP OF THE MODEL CALIBRATION RESPONSIVENESS SUMMARY TO THE MODELING STUDY FOR REST OF RIVER

The comments received from Reviewers on the Model Calibration Report will be used by EPA in conducting model validation, the third and final component of the Housatonic River modeling study, and in preparing the Model Validation Report, which will also be subject to Peer Review. Because the Peer Review comments will be addressed in the subsequent stages of the modeling study, EPA does not believe it is necessary or appropriate to issue a revised Model Calibration Report.

In conclusion, EPA recognizes the hard work and thought that the Reviewers contributed in conducting the Peer Review. Although EPA agrees with many of the comments provided by the Reviewers, EPA does not agree with some of the comments; these are documented in the responses and, in such cases, the technical basis for EPA's position is provided. EPA appreciates the effort from the Reviewers in providing their insights and believes that the modeling study, and in particular the model validation and report, will benefit greatly from their comments.

RESPONSE TO GENERAL ISSUES

1. GOALS OF THE MODELING STUDY

SUMMARY OF ISSUE:

Two Peer Reviewers indicated in their comments that to evaluate the adequacy of model calibration, it would have been helpful to have a better understanding of how the model will be used in the evaluation of remedial alternatives. In particular, Reviewers noted in their comments, and also during discussions at the public meeting, that it is important to know if the model will be used in a relative manner to compare alternatives to each other, or in an absolute manner to test the effectiveness of remedial alternatives against some standard or benchmark.

RESPONSE:

Under the terms of the Consent Decree governing the work in the Rest of River, EPA is to develop a framework for the Housatonic River modeling study and have it undergo Peer Review at three points: completion of the Framework, Model Calibration, and Model Validation. Upon completion of the Peer Review for Model Validation, EPA is to provide the model and all inputs and outputs to GE for their use in evaluating alternatives in the Corrective Measures Study (CMS). Neither the Consent Decree nor the Reissued RCRA Permit (Permit) specify the alternatives to be considered; however, these documents specify that the alternatives will be proposed by GE after receipt of the model in their submittal of the Corrective Measures Study Proposal. Thus, it is premature to begin discussion of the specific alternatives that will be evaluated, other than baseline conditions.

That being the case, it is expected the alternatives for addressing PCBs in the river channel and floodplain that will be modeled will include not only baseline conditions (monitored natural recovery), but also various forms of active remediation. Among the possible active remediation scenarios to be considered are dredging/excavation, capping, and various combinations of these techniques.

Because no reliable long-term estimates of future boundary conditions are available, the use of the model to predict future concentrations of contaminants necessarily must be based on a projection of the boundary conditions (flow, solids, and PCBs) over a period of decades. Because such projections will have an unknown degree of uncertainty associated with them that will impact model predictions, predictions of absolute concentrations are not anticipated to be accurate. Therefore, EPA will focus primarily on comparisons of relative performance among remedial alternatives against baseline conditions. However, the model predictions, in spite of the uncertainties noted above, will be the best estimate available of the potential magnitude of the expected reductions in exposure and will provide useful information in evaluating the performance of remedial alternatives.

The Interim Media Protection Goals (IMPGs), after approval by EPA, will be used as preliminary goals in evaluating the performance of the alternatives along with other factors specified in the Permit. The cleanup standards will be established by EPA in the Statement of Basis. Therefore, although there are no absolute criteria that will be used to evaluate model output, the predictions of PCB concentrations, combined with the results of the uncertainty evaluation, will be used, in a general way, to assess the effectiveness of alternatives to achieve target levels.

Because the assumption in the risk assessments that receptors average their exposure across a particular media is carried forward into the application of the IMPGs, model predictions of interest are the average concentrations of PCBs in tissue of aquatic receptors, and in water, river sediment, and floodplain soil. Although it is recognized that there is substantial variability in individual sediment and fish concentrations within each river reach, reproduction of this variability in model output is not necessary to achieve the goals of the modeling study; therefore, EFDC and FCM were not developed to simulate the distributions of individual PCB concentrations in the media of interest.

2. APPROACH TO CALIBRATION AND VALIDATION

SUMMARY OF ISSUE:

A majority of the Reviewers commented that the time period of 14 months over which EFDC was calibrated was too short to reflect many transport processes that may be significant over a longer duration. One Reviewer suggested a calibration period of 5 years, while another suggested 10 years. One Reviewer commented that the modeling strategy relied too heavily on model calibration.

In addition, some Reviewers suggested that some years of data be set aside for a "true" validation simulation, with no adjustments to the model parameterization. One Reviewer suggested an approach in which some of the originally planned validation period could be used to improve the calibration, with the remainder reserved for a true validation.

RESPONSE:

EPA agrees with the Reviewers and recognizes the importance of evaluating the performance of EFDC over periods longer than 14 months. The rationale for selecting the calibration period for EFDC was summarized in the Responsiveness Summary to the Peer Review of the Modeling Framework Design (MFD) and Quality Assurance Project Plan (QAPP) (WESTON, 2002):

> "The 1999 to 2000 calibration period was selected because this period coincides with the most recent, detailed data set collected by EPA,

> including the data obtained under storm event conditions. In selecting this 1-year period, which was characterized by a fairly wide range of flow conditions, the strategy is to first perform preliminary calibrations of the hydrodynamic, sediment transport, and PCB fate models under the higher flow (out-of-bank) storm event conditions. Next, the model calibration process will focus on the base flow conditions. Finally, because data were not collected for an event greater than 1.5 years during the calibration period, EPA will compare the model simulations to observations for two large storm events that occurred outside the calibration period (see Section 11, Rare Flood Events).
>
> "EPA believes that the advantages of using the high quality/intensity data set for model calibration outweigh the fact that the period is too short to see evidence of natural recovery, which has not been observed over the entire period of record in the PSA, and that a properly calibrated model will reliably represent conditions on a decadal scale. The 20-year period identified for model validation (1979-2000) is sufficient to demonstrate the ability of the model to simulate processes occurring on decadal time scales (see Section 22, Validation)."

Subsequently, the issue of the lack of change in PCB concentrations over the 14-month calibration period was discussed in Appendix B, Section B.1.1, of the Model Calibration Report:

> "Because the time scale of significant changes in PCB concentrations in the sediment bed is typically on the order of years, [model validation] will provide a test of processes that could not be evaluated completely in the 14-month calibration period. It is anticipated that adjustments to some calibration parameters will be required as part of the long-term validation phase. Subsequent to model validation, the models will be used to determine baseline conditions and to evaluate the effectiveness of remediation alternatives for the PSA."

It is also noted in the Model Calibration Report that longer-term simulations will provide a more robust or complete test of sedimentation rates and changes in sediment PCB concentrations simulated by the model.

Although the duration of the calibration period received little attention in the Reviewers' comments at the time of the Peer Review of the MFD (WESTON, 2000a) and the QAPP (WESTON, 2000b), this topic received attention from a majority of the Reviewers during the Model Calibration Peer Review. In response to these comments, EPA has modified the approach for the long-term simulations. EPA has reallocated the time periods used for calibration and validation by using the period from 1990 through June 2000 as a Phase 2 Calibration. This time period is consistent with the recommendations of several Reviewers, with suggestions that the longer calibration period range from 5 to 10 years. The Phase 2 Calibration simulations may result in adjustment to some model parameters, as expected in a model calibration exercise.

After completion of the Phase 2 Calibration, a continuous model simulation of the period 1979 through 2004 will be executed with no adjustments to model parameters. Model-data comparisons for the periods from 1979 through 1989, and July 2000 through 2004 will be presented as demonstrations of this model validation. With this approach, the model-data comparisons from the calibration period will be distinct from the model validation. In addition, EPA has extended the model domain downstream to include Reaches 7 and 8 for inclusion in the validation runs.

References:

WESTON (Weston Solutions, Inc.). 2000a. *Modeling Framework Design: Modeling Study of PCB Contamination in the Housatonic River.* Prepared for U.S. Army Corps of Engineers and U.S. Environmental Protection Agency. DCN GE-100500-AADX.

WESTON (Weston Solutions, Inc.). 2000b. *Quality Assurance Project Plan: Modeling Study of PCB Contamination in the Housatonic River.* Prepared for U.S. Army Corps of Engineers and U.S. Environmental Protection Agency. DCN GE-100500-AADY.

WESTON (Weston Solutions, Inc.). 2002. *Responsiveness Summary to the Peer Review of the Modeling Framework Design and Quality Assurance Project Plan: Modeling Study of PCB Contamination in the Housatonic River.* Prepared for U.S. Army Corps of Engineers and U.S. Environmental Protection Agency.

3. MODEL CONSTRUCT

SUMMARY OF ISSUE:

Three Reviewers commented on the structure of the linked models, primarily with the limitation that the computational requirements of EFDC place on addressing other desirable model questions. One Reviewer noted that the time necessary for each model run impacts the number of runs that can be used to explore model sensitivity and uncertainty. Another Reviewer recommended that a three-dimensional model be used for the river and a two-dimensional model be used for the floodplain. All three of these Reviewers suggested ways in which the computational burden could be reduced by modifying the way in which certain components of the models are linked and/or making changes to the model grid.

RESPONSE:

Innovative methods for reducing computational time have been considered over the course of model development. Several of these innovative approaches were implemented, while others were judged to be either ineffective or unnecessarily complex.

Ways to reduce computational time that were implemented during the model development effort included use of both a dynamic time step and a split time step; bypassing sections of model grid that are only changing slowly, if at all, during discrete periods of time; and conducting simulations for selected portions of the calibration period (e.g., high-flow events).

Unlike many models, which require a fixed time step during the entire simulation, EFDC allows the use of a dynamic time step; longer time steps were used during low-flow periods and shorter time steps were used during the less-frequent high-flow periods. EFDC also allowed implementation of a “split” time step (calculating derivatives in sediment cells less often than in the water column), which is appropriate because processes in the sediment generally occur on longer time scales than those in the water column.

Large portions of the model grid were bypassed during most of the simulation period, when flows were within the riverbanks. A majority of the grid cells in the model domain represent the floodplain, and it is only necessary to perform calculations in these cells during the infrequent periods when the floodplain is inundated.

Simulations for monitored storm events during which high-frequency data were collected were performed to evaluate the effect of model parameterizations without incurring the computational time necessary to conduct the full time-variable simulations. These selected periods were eventually included within a longer continuous simulation to eliminate potential artifacts that were introduced by assigning initial conditions in the sediment for the individual storm events.

Additional modifications that were examined and implemented included optimization of solution schemes and benchmarking performance of the model with a range of FORTRAN compilers and compiler options to determine the best combination for model simulations.

Ways to reduce computational time that were evaluated but not implemented included separating the three submodels of EFDC, use of a hybrid grid, flow binning, and parallel processing.

Separating the EFDC hydrodynamic simulation from the sediment transport and PCB fate simulations could be useful, but this would also decouple the feedback between the hydrology and sediment transport models in each time step, which was considered to be an important feature of a model of the Housatonic River system given the morphological changes in the river that have been observed.

Use of a model grid that differs substantially with respect to the degree of resolution and dimensionality in the main channel and the floodplain (three-dimensional versus two-dimensional, respectively) would make it more difficult to achieve conservation of momentum. The severity of this problem depends on how finely the bathymetry is represented with the variable grid resolution that is assigned to the channel and to the floodplain. The implementation of this approach for the Housatonic River would be extremely difficult, given the sharp meanders and variations in channel bathymetry.

Further, if the recommendations of the Reviewers were implemented it would also be necessary to incorporate a higher-resolution three-dimensional grid in the channel. This modification by itself would increase model simulation time and it is not clear if the net effect of the changes would, in fact, be a reduction in simulation time (see response to General Issue 4).

Use of average concentrations and/or flows to represent discrete flow intervals is another approach that has been used in other modeling studies with varying success. The effectiveness of this approach, commonly referred to as “binning,” depends on how coarsely the model inputs are binned. Use of steady-state flow conditions to span extended periods of time will reduce run time, but could also introduce a cumulative error that could become significant over a long-term simulation. There is a tradeoff between accuracy and the degree of binning that can be implemented, with greater improvements in simulation time but decreased accuracy as the number of discrete flows (i.e., flow bins) used to represent the hydrograph is decreased. Use of a series of steady-state flow bins would also produce a reduction in the variability of results simulated by the model, another issue that Reviewers commented on.

The approaches recommended by the Reviewers discussed above would also impact the evaluation of uncertainty, which the Reviewers indicated to be of importance. Analyses to evaluate some aspects of uncertainty would become more feasible with reduced computational requirements, but the uncertainty associated with the approximations could not be assessed readily.

The parallelization of the EFDC code was also evaluated as a means of reducing run time. However, it was concluded that the benefit of running EFDC on multiple processors was less than could be achieved by running the entire code on separate processors simultaneously.

4. NUMBER OF GRID CELLS ACROSS CHANNEL

SUMMARY OF ISSUE:

All seven Reviewers commented that the representation of the main river channel with a width of one cell in the EFDC model grid was problematical and did not allow for accurate simulation of the lateral variation for certain important processes and variables. Processes and parameters noted most often in this regard included flow, sediment-water contaminant flux, bottom shear stress, and erosion potential. One Reviewer noted that the necessary lateral averaging of parameters used to evaluate the calibration (e.g., TSS) would have the effect of artificially improving the apparent goodness of calibration. Most Reviewers recommended a minimum of three lateral cells to represent the channel.

In general, the Reviewers recognized that the use of multiple lateral cells to represent the channel may lead to computational difficulties, and in some cases, suggested ways in which the computational burden of additional channel cells might be reduced.

RESPONSE:

The complexity of the Primary Study Area (PSA) poses a challenge for developing a computationally efficient but physically comprehensive model. Due to the complex bathymetry and morphology of the Housatonic River, the circulation and flow regimes are also complex.

At the time of the MFD Peer Review, EPA proposed the following regarding the EFDC grid:

> "The primary focus of the grid development effort proposed in the MFD is on the use of a curvilinear grid for primary channel/proximal floodplain (called the "main channel") and a separate linked grid which represents the distal floodplain (see the following Figure, where the inner three cells represent the river channel, the outer two cells represent the proximal floodplain, connected to the distal floodplain cells)." (EPA Response to Peer Review Panelist Questions on the Housatonic River Modeling Framework Design, 4/12/2001, pp. 54-55).

Several Reviewers criticized this representation of the river by three grid cells across the channel width. One Reviewer discussed the importance of secondary flows, implying that he thought a three-dimensional representation of the river channel was required, and recommended against the use of the computational grid shown in Figure GI-4-1. Another Reviewer commented that a three-dimensional grid was not necessary, but recommended additional resolution in the river channel that would result in 4 to 5 cells across the river in much of the model domain. Other Reviewers believed that simpler approaches would be better.

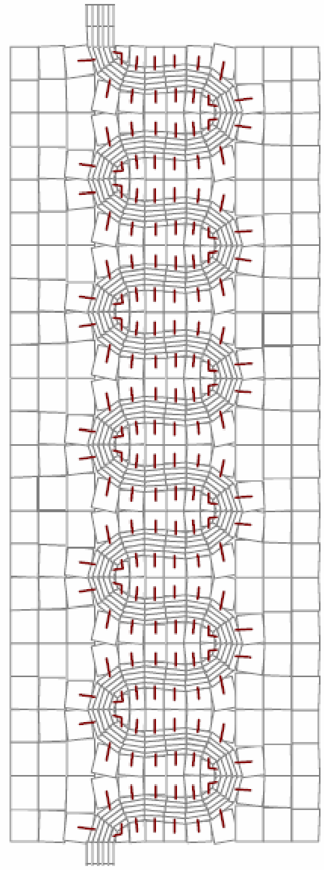


Figure GI-4-1 Grid Proposed by EPA During MFD Peer Review

It is well known that flow in a river, especially in the vicinity of a river bend, is fundamentally a three-dimensional process. Rozovskii (1957) conducted one of the first hydraulic experiments of secondary circulation created by a steady, open channel flow around a bend. In his research he showed that as the river flow enters the bend, it veers toward the outside bank due to centrifugal acceleration. This, in turn, piles up water at the outer wall of the bend, creating a water level slope in the cross-channel direction, a phenomenon known as superelevation. A secondary (lateral) circulation cell is thus established by the three-dimensional balance of centrifugal acceleration, barotropic pressure gradient, and vertical shear stress; water flows toward the outer bank at the surface, downwells, flows back toward the inner bank at the bottom, and upwells there. Superimposed on this secondary circulation is the primary movement of water downstream, completing the phenomenon known as helical motion. Boxall et al. (2003) conducted hydraulic experiments to evaluate the effect of channel curvature on the transverse mixing coefficient of a meandering open channel, concluding that the transverse mixing coefficient varies in direct relation to channel curvature and that the variation is cyclic with geometry: it increases at the apex of the bend, and decreases in straighter regions.

However, recognizing the practicalities associated with modeling large, complex river systems, the state-of-the-art in river modeling (Choi and Kang, 2004; Makhanov et al., 1999; Yoon and Kang, 2004) is to use vertically integrated model physics (one layer), which is the approach used in the Housatonic River modeling study. Although previous studies use this approach, it is recognized that the limitations on the representation of the physics impose certain constraints.

To properly represent the cross-section of a river, on the order of 10 grid cells are required (Alan Blumberg, personal communication, 2005). The use of three boxes may seem to provide a way to include lateral structure in the flow field, but three boxes in a numerical model cannot simulate any of the lateral dispersion because at least five boxes are required to represent the second derivative of the velocity field (Alan Blumberg, personal communication, 2005). In addition, incorporating lateral segmentation in the channel would increase the time required to perform multi-decadal simulations to several months. Moreover, it makes little sense to dramatically increase the lateral grid resolution without also using three-dimensional model physics, which is infeasible because of computational resource constraints.

Several Reviewers, in their comments on the MFD, recognized that a three-dimensional model was neither feasible nor necessary, and in several cases recommended that EPA consider a one-dimensional model:

> "A three-dimensional, time-dependent model consumes much more development and computational time and is probably no more accurate in practice than a two dimensional, time-dependent model (with a correction for quasi-equilibrium distribution of sediment in the vertical). This latter model is also much more computationally efficient. This has been shown in numerous cases. Even for the pond, a two-dimensional model is sufficiently accurate to predict sediment and PCB transport." (W. Lick, May 17, 2001)

> "It is likely that a useful model could be developed that was restricted to in-bank flow. It is also highly likely that modeling of the significant over bank flows could also be successfully completed. However, it would require a different model in each case. In fact, one-dimensional models are widely used for these purposes. Specific examples include HEC-2 (Corps of Engineers), NWS Flood Wave, Fischer Delta Model (Hugo B. Fischer, Inc.), DWRDSM2 (Calif. Dept. of Water Resources), and there are probably many others." (J. List, May 22, 2001)

Once the decision was made in response to the Reviewers' comments on the MFD not to use a three-dimensional segmentation of the water column, much of the ability to explicitly account for the lateral processes was eliminated. Therefore, the use of a vertically integrated equation set introduces the need to use parameterizations in place of an explicit representation.

A practical, numerically efficient, yet accurate approach using a variable-resolution orthogonal curvilinear grid was taken to discretize the entire PSA. The grid is designed to follow the main channel as much as possible. The typical grid spacing is on the order of 20 meters (m). The use of a grid resolution of 20 m within the main channel results in nearly the entire channel width being covered by a single cell. This coarse segmentation does not provide the ability to simulate the lateral variations in flow known to exist in a meandering channel. However, the grid is able to incorporate the floodplain, main channel, adjacent backwaters, and the complex oxbows near the headwaters of Woods Pond as a contiguous system, responding to Reviewers' concerns about the "nested grid" proposed in the MFD.

Decisions on grid resolution need to balance the level of physics the model is capable of simulating and that is needed to achieve the goals of the study, computational resources, and other considerations. EPA believes that the current curvilinear grid achieves such a balance, including the extent of coverage of the PSA by the grid, the ability of the grid to adequately represent the river channel, and the use of vertically integrated physics.

References:

Blumberg, A. 2005. Personal communication (e-mail) from Alan Blumberg, Stevens Institute of Technology, to Edward Garland, HydroQual, re: grid issues.

Boxall, J.B., I. Guymer, and A. Marion. 2003. Transverse mixing in sinuous natural open channel flows. *J. Hydraulic Res.* 41:153-165.

Choi, S. and H. Kang, 2004. Depth-averaged modeling of flows and sediment transport in open-channel with submerged vegetation. ASCE Water World Congress.

EPA (U.S. Environmental Protection Agency). 2001. EPA Response to the Peer Review Panelist Questions on the Housatonic River Modeling Framework Design. April 2001.

Makhanov, S., S. Vannakrairojn, and E. Vanderperre, 1999. 2D numerical model of flooding in East Bangkok. ASCE, *J. Hydraulic Engineering*, Vol. 125, No. 4, pp. 407-414. April 1999.

Rozovskii, I.L. 1957. Flow of water in bends of open channels. Academy of Sciences of the Ukrainian S.S.R. Institute of Hydrology and Hydraulic Engineering, Kiev. Translated from Russian by Y. Prushansky, Israel Program for Science Translation, S. Monson, Jerusalem. PST Catalog No. 363. 233 pp.

Yoon, T., and S. Kang, 2004. Finite volume model for two-dimensional shallow water flows on unstructured grids. ASCE, *J. Hydraulic Engineering*, Vol. 130, No. 7, pp. 678-688. July 2004.

5. PCB FATE

SUMMARY OF ISSUE:

Four Reviewers commented on the high degree of small-scale spatial variability in PCB concentrations in main channel sediment, but differed on the implications of the variability for achieving the goals of the modeling effort. One Reviewer indicated that the inability of the model to reproduce variability on such a small scale is to be expected and does not constitute a problem for the modeling. Another Reviewer simply indicated that the spatial variability was "perplexing" because of the lack of apparent temporal variability. A third Reviewer, however, saw both the lack of knowledge concerning the process or processes that led to such variability, and the inability of the model to reproduce the variability, as a major issue that makes the modeling approach followed thus far "inappropriate."

Five of the seven Reviewers recommended that volatilization be considered for inclusion as a loss process in the contaminant transport portion of the model, with most indicating that they believed volatilization could be significant and therefore, should be included. One Reviewer noted that any calculations done to estimate potential loss due to volatilization should be based on more current values for Henry's Law constant than were used in the Resource Conservation and Recovery Act Facility Investigation (RFI). Two Reviewers noted that volatilization loss would be particularly important from the floodplain because of the relatively greater exposure area and solar heating.

One Reviewer recommended that the modeling effort focus on modeling of individual PCB congeners, rather than the current approach of modeling total PCBs (tPCBs) within EFDC and adding the modeling of selected coplanar congeners in FCM. This Reviewer noted that modeling of tPCBs is not the "state-of-the-art in modeling methodology" but also recognized that computational considerations for EFDC would make such an approach difficult. The Reviewer also provided an approach to evaluating model bias that makes use of congener-specific model results, and further recommended that congener-specific degradation rates be included in the modeling.

RESPONSE:

Small-Scale Spatial Variability and Equilibrium Partitioning

EPA is also not satisfied with the lack of a mechanistic explanation of PCB fate processes for the small-scale variability, and spent considerable effort attempting to develop such an explanation. However, EPA does not agree with the Reviewers who believed that the lack of such an explanation compromises the ability of the model to achieve the goals of the modeling study. In addition, EPA does not agree that there is any inconsistency between the presence of pronounced small-scale variability and the inability to resolve temporal variability in the data. The sampling design was not intended to provide data relative to short-term temporal variability and any such temporal variability that does exist would be masked by the spatial variability.

The small-scale variability noted by Reviewers was identified as part of the benthic community investigations, which involved collection of 12 replicate grab samples of sediment from a small area, in some cases less than 1 square meter (m^2). Results of PCB analyses of aliquots from these samples indicated variability in some locations of over two orders of magnitude in PCB concentrations, with no such differences apparent in other factors known to be correlated with contaminant concentration (e.g., sediment grain size, total organic carbon [TOC]). Although this high degree of variability was observed in eight of the nine benthic community sampling locations in the PSA, it is assumed to occur throughout the PSA because other efforts were made to resample specific locations with the result being highly variable PCB concentrations over small time scales as well.

In addition to the small-scale spatial variability in PCB concentrations, EPA also noted that a subset of the samples in Reach 5A did not follow traditional equilibrium partitioning (EqP) behavior. In these samples, PCB concentrations were elevated well above concentrations that would be predicted based on organic carbon content and, consequently, these samples appeared as outliers when the data were carbon-normalized. As part of the investigation of this latter phenomenon, microscopic examination of samples of Reach 5A sediment indicated the presence of a film or coating on individual quartz grains. Samples were sent to Sandia National Laboratories (SNL) and examined by scanning electron microscopy (SEM), X-ray diffraction (XRD), and energy-dispersive system (EDS) via microprobe. These studies indicated the presence of carbon and chlorine in the film, which was identified by the Sandia scientists as representing PCBs present in the film. This resulted in EPA's working hypothesis that PCBs in these samples are present as a recalcitrant coating on individual quartz particles, with consequent reduced bioavailability.

It is possible that the two phenomena, i.e., small-scale spatial variability and high PCB concentration surface films on individual sand particles, are related. Transport of the coated grains is a stochastic process that could result in greater numbers of these grains in some samples than others, even over very small spatial scales. The presence of a greater or lesser number of such particles could affect the overall PCB concentration in a sample without apparent differences in other sediment characteristics. This hypothesis would also explain the observation that some sediment does not conform to EqP theory because of the presence of greater amounts of coated sands. The greatest amount of small-scale spatial variability in PCB concentrations was observed in Reach 5A, where higher energy hydrodynamics result in greater amounts of coarse sand in bed sediment. This hypothesis, however, would not appear to explain the variability also observed at benthic community sampling locations in Reach 5C, where relatively small fractions of the sediment are sand, unless a similar phenomenon is occurring on the cohesive size fractions, which is not known.

Regardless of the explanation(s), or lack thereof, for these phenomena, EPA believes that it is not necessary to understand the processes that underlie them, or to include such processes, or the resultant variability, in the modeling framework. With regard to small-scale variability in PCB concentrations, the variability occurs on a spatial scale that is far smaller than the cells of the modeling grid; therefore, the model grid is unable to resolve these differences or to make use of any process-oriented explanation that might be developed. Because both human and ecological receptors are exposed to the range of concentrations in an area rather than only to the concentration in an area equivalent to a benthic grab sample, they act as integrators of the variability in concentrations and the exposure point concentration (EPC) of interest is larger than an individual model grid cell. Transport processes are simulated on a scale that affects the bed sediment in each model cell equally, so again it is not necessary to explicitly account for the small-scale variability.

In the case of sediment that does not conform to EqP theory, the only component of the modeling framework that is potentially affected is the bioaccumulation model. The issue of sediment that does not conform to EqP theory was analyzed and discussed at length in the Model Calibration Report (Section C.3.1.2) (WESTON, 2004b) and a procedure was developed to arrive at adequate calibration. EPA believes that the proposed approach will achieve the goals of the modeling study.

Volatilization

In the October 2000 Modeling Framework Design (MFD) document that was reviewed by the Peer Review Panel in 2002, volatilization was described as a potentially important process that was still being evaluated for inclusion in the modeling framework (Section 3.3.4.2). No comments were received from the Reviewers on the inclusion or exclusion of volatilization as a process in the model during Peer Review of that document.

As discussed in Section 4.5.5 of the final MFD (WESTON, 2004a), an analysis of the potential importance of volatilization was conducted by BBL and QEA and presented in the RFI Report. Using conservative assumptions to derive an upper-bound estimate, and using a Henry's Law constant derived from published values (Brunner et al., 1990), that analysis concluded that the loss of water column PCBs via volatilization in Woods Pond under low-flow conditions would be only 5%. Accordingly, EPA indicated in the final MFD that the process of volatilization was not sufficiently significant to be included in the modeling study.

In response to comments received from Reviewers regarding the decision to exclude volatilization as a PCB fate process, and particularly in response to a comment from one Reviewer providing a reference to more recent Henry's Law constants for PCBs (Bamford et al., 2000), EPA has reanalyzed the potential loss of PCBs from the PSA via volatilization. The reanalysis was conducted based on congener-specific Henry's Law constants and the relative congener composition of the predominant Aroclors present in the PSA (A1260 and A1254). The results of the reanalysis indicate that volatilization loss of PCBs from the main channel reaches is likely still an insignificant fate process; however, the increased residence time of contaminated surface water in backwaters could result in sufficient loss of PCBs to justify inclusion of volatilization in the modeling framework. Accordingly, EPA will now include volatilization as a loss process in the contaminant transport component of the model, and has modified the EFDC code to activate the volatilization subroutine.

EPA does not believe, however, that it is necessary or advisable to include volatilization from the floodplain as a loss process. The concentration of PCBs present in the vapor phase in the soil is extremely low, even considering the updated Henry's Law constants; therefore, volatilization loss of PCBs from the floodplain does not represent a significant PCB fate process in the PSA. Inclusion of this process in the floodplain would require each floodplain cell to be evaluated at every time step, rather than the current practice of performing calculations in floodplain cells only when the floodplain is inundated. This would represent a considerable computational burden and would compromise many other aspects of the modeling study, without commensurate gain, particularly considering the substantial uncertainty involved in parameterizing this process.

Modeling of PCB Congeners

As discussed in Section 4.5.4 of the final MFD, EPA recognizes that individual PCB congeners exhibit different behavior with respect to chemical, physical, and biological processes that control transport and fate, and agrees that modeling of individual congeners could be included in a "state-of-the-art" modeling study. It is important, however, to evaluate the efficacy of the current modeling framework and model calibration in the context of the goals of the modeling study and the intended use of the model. As discussed in the response to General Issue 1, the Housatonic River model is intended to be used to differentiate between the effectiveness of a limited number of remedial alternatives, and to estimate the time necessary for tPCBs, which is the primary measure of PCB concentrations of interest in the risk assessments, in fish tissue and other media to reach certain concentrations under those alternatives. With that overall objective in mind, the question of whether it is necessary to model individual congeners can be addressed in two forms:

- Is it necessary to have information on congener concentrations to distinguish between the relative effectiveness of remedial alternatives?

- Is it necessary to model individual congeners to arrive at a simulated concentration of tPCBs that can be used to compare relative effectiveness of remedial alternatives, or to compare to a standard or other established concentration?

EPA believes it is clear that the response to the first question is that it is not necessary. Whatever processes may be operating differentially on individual congeners, and whatever errors, if any, may be introduced by simplifying the complex behavior of a mixture of PCBs into “average” properties for the mixture, would operate similarly for the various remedial alternatives tested via the model. Consequently, modeling tPCBs will satisfy the first overall goal of the modeling, that of comparing the relative effectiveness of remedial alternatives.

The second question will be addressed by the results of the model calibration and validation, which is being conducted on tPCBs (except for the modeling of a few selected congeners as a test of the robustness of the bioaccumulation model). If the calibration and validation efforts can achieve adequate agreement with data using the tPCB approach, that outcome will be sufficient demonstration that it is not necessary to model individual congeners. Based on the results seen to date, EPA believes that the final results of the ongoing calibration and validation will provide just such a demonstration.

In addition, as discussed in detail in Section 3.5 of the final MFD, PCB congener composition is relatively stable throughout the various reaches and media in the PSA and downstream. Because of this relative stability, it is possible to use observed ratios of concentrations of individual congeners to tPCBs to calculate probable concentrations for selected congeners from data on tPCBs. This approach was used, for example, in the Human Health Risk Assessment to evaluate risk due to coplanar PCB congeners that contribute to TEQ risk. Therefore, the fact that individual congeners are not modeled explicitly does not preclude applying the results of the modeling study to specific questions involving individual congeners.

References:

Bamford, H., D. Poster, and J. Baker. 2000. Henry’s Law constants of polychlorinated biphenyl congeners and their variation with temperature. *J. Chem. Eng. Data* 45:1069-1074.

Brunner, S., E. Hornung, H. Santi, E. Wolff, O.G. Piringer, J. Altschuh, and R. Brüggemann. 1990. Henry’s Law constants for polychlorinated biphenyls: Experimental determination and structure-property relationships. *Environ. Sci. Technology* 24:1751-1754.

WESTON (Weston Solutions, Inc.). 2004a. *Modeling Framework Design: Modeling Study of PCB Contamination in the Housatonic River*. Prepared for U.S. Army Corps of Engineers and U.S. Environmental Protection Agency. DCN GE-042104-ACDP.

WESTON (Weston Solutions, Inc.). 2004b. *Model Calibration: Modeling Study of PCB Contamination in the Housatonic River*. Prepared for U.S. Army Corps of

Engineers and U.S. Environmental Protection Agency. DCN GE-122304-ACMG.

6. DEPTH OF SEDIMENT BIOAVAILABLE LAYER

SUMMARY OF ISSUE:

Three Reviewers indicated that the 6-inch active (mixed or bioavailable) layer depth used in the model during the Calibration was too thick. These Reviewers did not agree on an alternative depth for this layer, and suggested values that ranged from 3 inches (7.5 cm) to as little as approximately 1 inch (2.5 cm). All three Reviewers agreed that the value assigned as the depth of this layer would have a significant effect on model results, and particularly on the estimated time necessary for natural recovery of the system.

RESPONSE:

EPA agrees that the depth of the bioavailable sediment layer is important for the prediction of long-term PCB fate. In Attachment C.19 of the Model Calibration Report, rationales were provided for the use of a 6-inch surface layer in providing unbiased estimates of a biologically relevant exposure depth over the calibration period. For longer-term simulations, a better estimate of the well-mixed sediment layer was required. It was acknowledged in Attachment C.19 that a 6-inch (15-cm) layer may "overestimate the depth to which most bioturbation occurs in Housatonic River sediment."

An extensive literature review has been conducted to refine the estimates of the depth of biological mixing in sediment. The methods and results of the literature review are provided in Attachment 1 (Freshwater Bioturbation Depth) to this Responsiveness Summary. A summary of the main findings is provided in Table GI-6-1. Based on the review of the literature on freshwater bioturbation, and considering site-specific data on habitat and resident benthic communities, bioturbation depths were estimated for each of three PSA habitat types. The biologically mixed depth, which represents the sediment layer that is thoroughly mixed due to bioturbation, is estimated to range from approximately 4 cm to approximately 10 cm across the PSA, depending on habitat. Deeper, but less-pronounced, biological mixing is estimated to occur at depths up to 20 cm below the sediment-water interface (i.e., biologically influenced depth). The ranges of biologically mixed depths provided in Table GI-6-1 are similar to the estimates provided by the Reviewers and also by GE in the comments on the Model Calibration Report (QEA, 2005).

In the application to long-term model simulations, the range of values for bioturbation depths in Table GI-6-1 will be used as one factor to guide the specification of biologically mixed depth and biologically influenced depth. The selected values will be combined with mixing coefficients that reflect the differences in magnitude of mixing within these surface sediment layers and the final parameterization will be compared to the site-specific data to ensure reasonableness. Additional abiotic sediment mixing (i.e., erosion and deposition) is handled separately within EFDC. For the EFDC to FCM model linkage, the bioavailable depth will be defined as computationally equivalent to the biologically mixed depth.

Table GI-6-1

Summary of Biological Mixing Depths Applicable to the PSA

Habitat		Biologically Mixed Depth Interval *	Biologically Influenced Depth Interval *
1	Upstream main channel (Reach 5A)	0 – 4 cm	4 – 10 cm
2	Downstream (Reaches 5B, 5C, Woods Pond)	0 – 8 cm	8 – 15 cm
3	Backwaters (Reach 5D)	0 – 10 cm	10 – 20 cm

* The depths shown represent depth intervals below the sediment surface. The total bioturbation depth is represented by the bottom of the biologically influenced depth interval (e.g., 10 cm for Reach 5A). Definitions for terms are provided in Section 1.2 of Attachment 1.

Reference:

QEA (Quantitative Environmental Analysis, LLC). 2005. *Comments of the General Electric Company on the U.S. Environmental Protection Agency's Model Calibration: Modeling of PCB Contamination in the Housatonic River*. Prepared for General Electric Company, Pittsfield, MA by Quantitative Environmental Analysis, LLC, Liverpool, NY. February 7, 2005.

7. RESUSPENSION AND DEPOSITION

SUMMARY OF ISSUE:

Three Reviewers commented on the formulation and calibration of the processes of resuspension and deposition in the EFDC sediment transport model. These Reviewers indicated that processes involving multiple variables may be inadvertently calibrated incorrectly because offsetting errors in individual parameters may produce satisfactory agreement with the data to which the parameters are being calibrated. These Reviewers also expressed concern that incorrect values for the parameters of resuspension and deposition could nonetheless lead to correct prediction of total suspended solids (TSS).

In extensive comments regarding this issue, one Reviewer provided a detailed discussion of resuspension processes, indicating the exponent used in the model formulation for calculating resuspension on the basis of excess shear stress was incorrect and recommending additional Sedflume experiments. This Reviewer was also critical of the effect that the 6-inch sediment active layer would have on the process of bed armoring. The Reviewer also commented on the deposition formulation used in the EFDC model, indicating that a formulation for time-dependent flocculation is available and should be included.

RESPONSE:

There are several technical issues included within this General Issue; they are discussed individually below.

Role of Resuspension Flux in Overall Solids Balance

Numerous figures in the Model Calibration Report summarize the net effect of sediment resuspension and deposition processes. Evaluation of these figures led one Reviewer to question if the sediment bed of the PSA is a significant source of solids transported within the model domain. In Figure GI-7-1, the magnitudes of erosion and deposition are summarized by reach and compared to the solids entering the upstream boundary and exiting the downstream boundary. Because water column measurements at New Lenox Road reflect the effect of resuspension and deposition processes in Reach 5A and only a portion of Reach 5B, results for Reach 5B are separated into two subreaches, upstream and downstream of New Lenox Road. Results are summarized for two time periods, Event 1 (May 18-23, 1999), which was the largest monitored event, and the 14-month Phase 1 Calibration period.

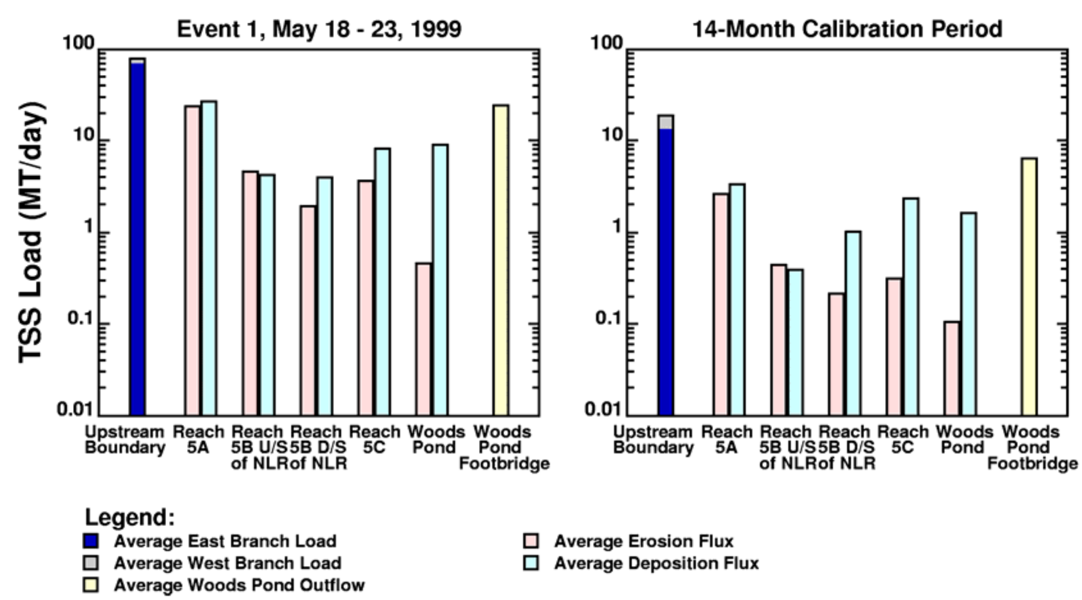


Figure GI-7-1 Erosion and Deposition Flux by Reach for Storm Event 1 and 14-Month Calibration Period

During the 5-day period associated with Event 1, the mass of solids resuspended from the main channel of the PSA was 43% of the mass of solids entering the upstream boundaries. Over the 14-month Phase 1 Calibration period, resuspension within the PSA was approximately 20% of the solids input at the upstream boundaries. Resuspension mass fluxes during the May 1999 high-flow event were approximately an order of magnitude greater than the averages over the Phase I Calibration period, a difference that is consistent with the conceptual model for sediment transport. A consistent pattern of decreasing magnitude of resuspension fluxes in the downstream direction was apparent over both periods due to the decreasing velocity resulting from the increase in cross-sectional area and backwater effects from Woods Pond Dam. Resuspension fluxes in the higher-energy environment of Reach 5A account for approximately 70% of the resuspension fluxes from the river channel over both time periods. These results are consistent with EPA's conceptual model, in which resuspension is an important process in the Housatonic River PSA.

Constraints on Resuspension and Deposition Rates

Data describing changes in channel cross section, radioisotope data (used to estimate net sedimentation rates), and water column suspended solids concentrations reflect the net effect of solids deposition and resuspension. Some Reviewers expressed concern that an infinite number of combinations of different resuspension and deposition rates could be specified to reproduce water column suspended solids, without any constraints on the pairs of rates selected. This concern does not recognize that there is a constraint imposed by PCB transport that results from resuspension and deposition processes.

Mr. Endicott recognized this point in his comments, noting:

> "In some other river systems, errors in deposition and resuspension fluxes are revealed during calibration of water column PCB concentrations, because the concentration gradient between sediment and water amplifies the error. In the PSA this will not work so well, because there is only a small gradient between suspended and bedded particulate PCB concentrations." (see Specific Response 1-DE-8).

It is acknowledged that the difference between water column and bed sediment particulate PCB concentrations in the Housatonic River may not be as large as in some other systems, when compared on a whole sediment basis. However, the difference in concentrations is amplified considerably when the PCBs associated with the fine sediment fraction are compared to the water column particulate PCBs. This is because the fine sediment fraction is higher in organic carbon content and thus, has a relatively high dry weight PCB concentration as well (see Calibration Report Figures B.4-22 through B.4-25).

Because the fine sediment fraction is the predominant fraction that contributes to the resuspension flux, the PCB concentration associated with the fines serves as a useful tracer that places a constraint on the flux of solids to the water column. The solids flux between the water column and sediment bed needs to be such that water column TSS levels are correctly simulated, although it is acknowledged that correct simulation of TSS can be achieved by a combination of incorrect settling and resuspension rates and therefore, is not, in itself, sufficient for demonstrating calibration. However, the spatial profile of water column PCB concentrations provides an additional constraint on the magnitude of the settling and resuspension fluxes of TSS to the water column. In combination, these comparisons (for both solids and PCBs) indicate the reasonableness of the calibrated model.

To demonstrate the constraints that water column PCB concentrations place on the specification of resuspension and settling rates, an alternate parameterization of resuspension and deposition was specified for a simulation of Event 1. The alternate parameterization consisted of a reduction in the critical shear stress for cohesive solids resuspension and an increase in the cohesive solids settling velocity, resulting in an increase in the exchange between water column and bed solids. Results from the original Phase 1 Calibration and this alternate parameterization of resuspension and deposition are presented in Figure GI-7-2. Although the increased resuspension flux of solids from the bed was offset by increased deposition, and similar water column TSS profiles were achieved, water column PCB concentrations increased by approximately a factor of five, making the simulation results from the alternate parameterization inconsistent with the data. Thus, Mr. Endicott's observation regarding the constraint that reproducing both suspended solids and water column PCB concentrations imposes on parameterization of resuspension and deposition applies to the Housatonic River, and the satisfactory reproduction of water column PCB concentrations demonstrates that resuspension and deposition processes have been parameterized correctly.

Non-Cohesive and Cohesive Transport Formulations

With regard to the questions raised regarding transport formulations, there are a number of formulations accepted by modelers in common practice representing transport in riverine systems, with no consensus on which formulation is most applicable to a particular system. There are five individual processes that, collectively, comprise the transport formulations:

- Deposition of non-cohesive solids.
- Deposition of cohesive solids.
- Erosion of non-cohesive solids.
- Resuspension of cohesive solids.
- Bedload transport of non-cohesive solids.

Deposition of non-cohesive solids is described by the formulations in van Rijn (1984), which are well established and have been used in numerous sediment transport models.

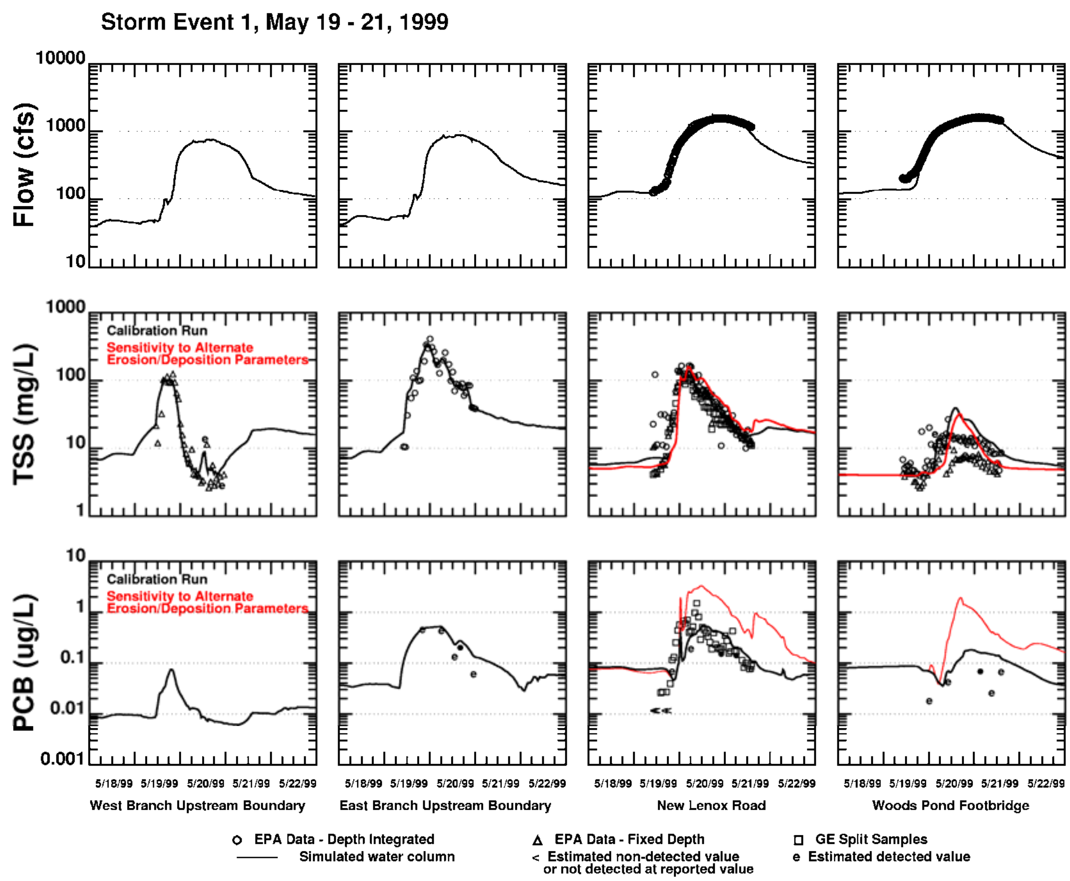


Figure GI-7-2 Concentrations of Total Suspended Solids (TSS) and PCBs for Original and Alternate Phase 1 Model Calibration

Deposition of cohesive solids is a complicated process that is affected by variations in, for example, mineralogy, grain size distribution, and suspended solids concentration. This process can be represented at different levels of complexity, ranging from constant settling velocity to complex flocculation models. The settling velocity used in EFDC is a function of the weighted average of the settling velocities for washload and suspended load. The EPA modeling team's efforts to implement Dr. Lick's simple flocculation model, in response to a recommendation provided during the Calibration Peer Review, are discussed below.

The equations used to describe the erosion of non-cohesive solids were developed by Garcia and Parker (1991) based on comparisons between simulations and experimental and field observations. Particle sorting and hiding mechanisms that are responsible for bed armoring for non-uniform sediment mixtures are included in the formulation.

Resuspension of cohesive solids is highly site-specific, and generic equations similar to those discussed above for erosion of non-cohesive solids are not available for cohesive solids. The parameterization of resuspension of cohesive solids used in EFDC was based on analysis of erosion data collected in the Sedflume experiments on Housatonic River sediment. This analysis was presented in Attachment B.5 of the Model Calibration Report.

The Sedflume data were also reanalyzed in accordance with the approach and formulations suggested by Dr. Lick in his review comments; however, the variability in the Housatonic River data was not explained by those equations and analytical approach. The results of the analysis are shown in Figures GI-7-3a, b, and c. Each figure displays the data for a different particle size (D_{50} < 30 μm, 30 μm ≤ D_{50} < 100 μm, and D_{50} ≥ 100 μm, respectively) to reduce the particle size effect, and each panel presents data for different depth intervals (<5 cm, 5–10 cm, 10-15 cm, etc.) in an effort to account for changes in bulk density. When stratified in this manner, there is still considerable scatter in the data, an indication that factors other than particle size and bulk density are affecting the magnitude of the shear stress needed to resuspend Housatonic River sediment.

The range of shear stresses applied in the Sedflume experiments conducted on Housatonic River sediment cores encompassed a high percentage of the range of shear stresses encountered in the PSA, and therefore, the formulations derived from the site-specific data do not need to be extrapolated substantially beyond the range of the data, even in application to extreme flow conditions. Differences between the exponent of the cohesive solids erosion formulation derived from analysis of site-specific Sedflume data and the central tendency of values derived from Sedflume experiments on other systems do not contribute to substantial differences in cohesive solids erosion because of the relatively small range of extrapolation beyond the range of the site-specific data.

Non-cohesive bedload transport was described by the modified Engelund-Hansen formulation (Wu et al., 2000), which has been used in other sediment transport models. The Engelund-Hansen formulation was selected from the formulations available within EFDC because it had a number of desirable characteristics, including the ability to represent the effect of variation in sediment grain size within an individual grid cell, and the parameterization could be assigned from available site-specific information. The formulation also produced reasonable agreement with site-specific bedload data (Section B.3.2.3.2; Figure B.3-22) collected as part of EPA's monitoring efforts to support the modeling study.

EPA believes that the model, as modified in response to the Peer Reviewers' comments, is able to adequately simulate sediment/contaminant transport in the study area and will provide a useful tool to evaluate the relative performance of potential remedial alternatives.

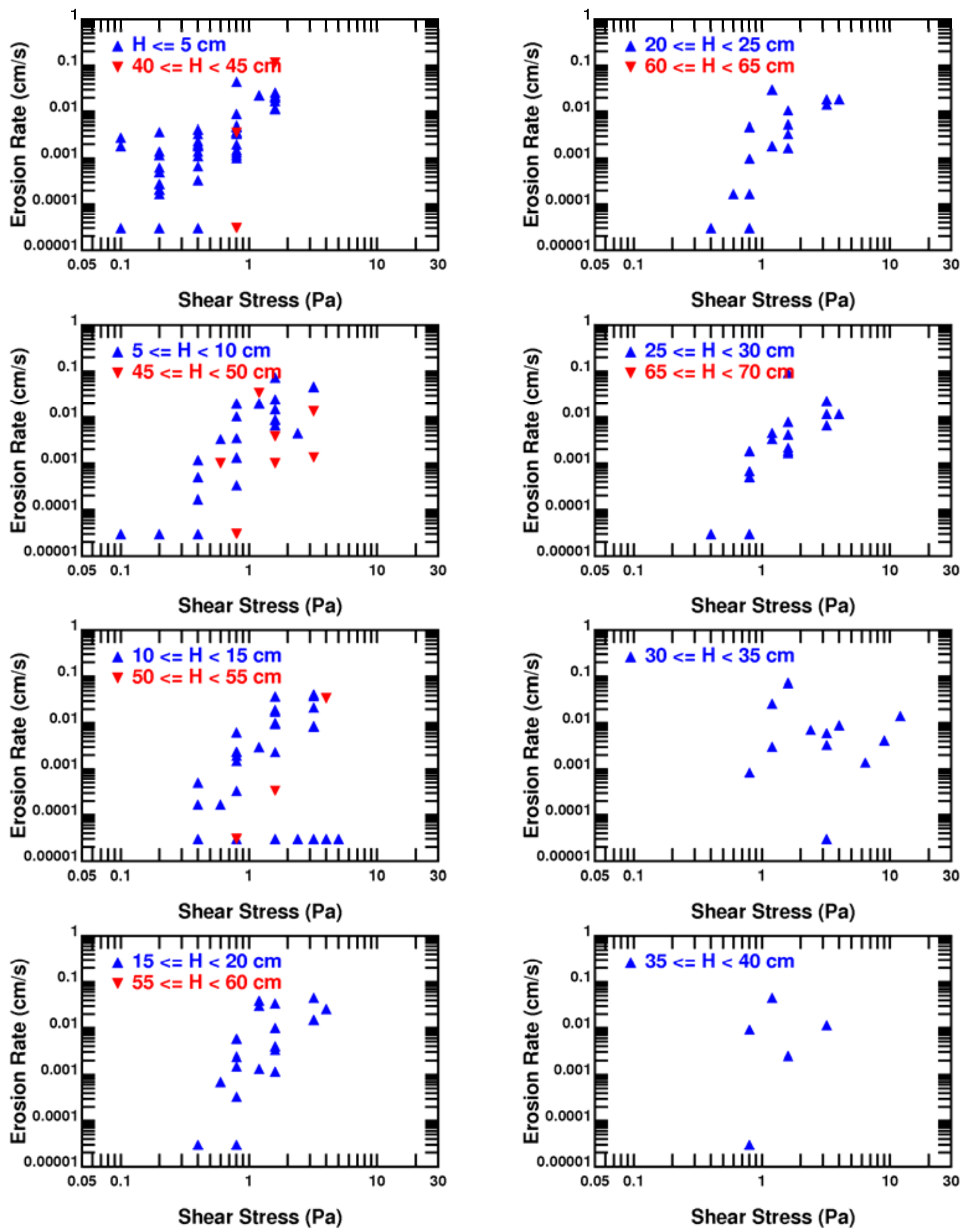


Figure GI-7-3a Erosion Rate Versus Shear Stress (as a function of depth in core) for Sedflume Experiments on Sediment with Median Particle Diameter (D_{50}) Less than 30 μm

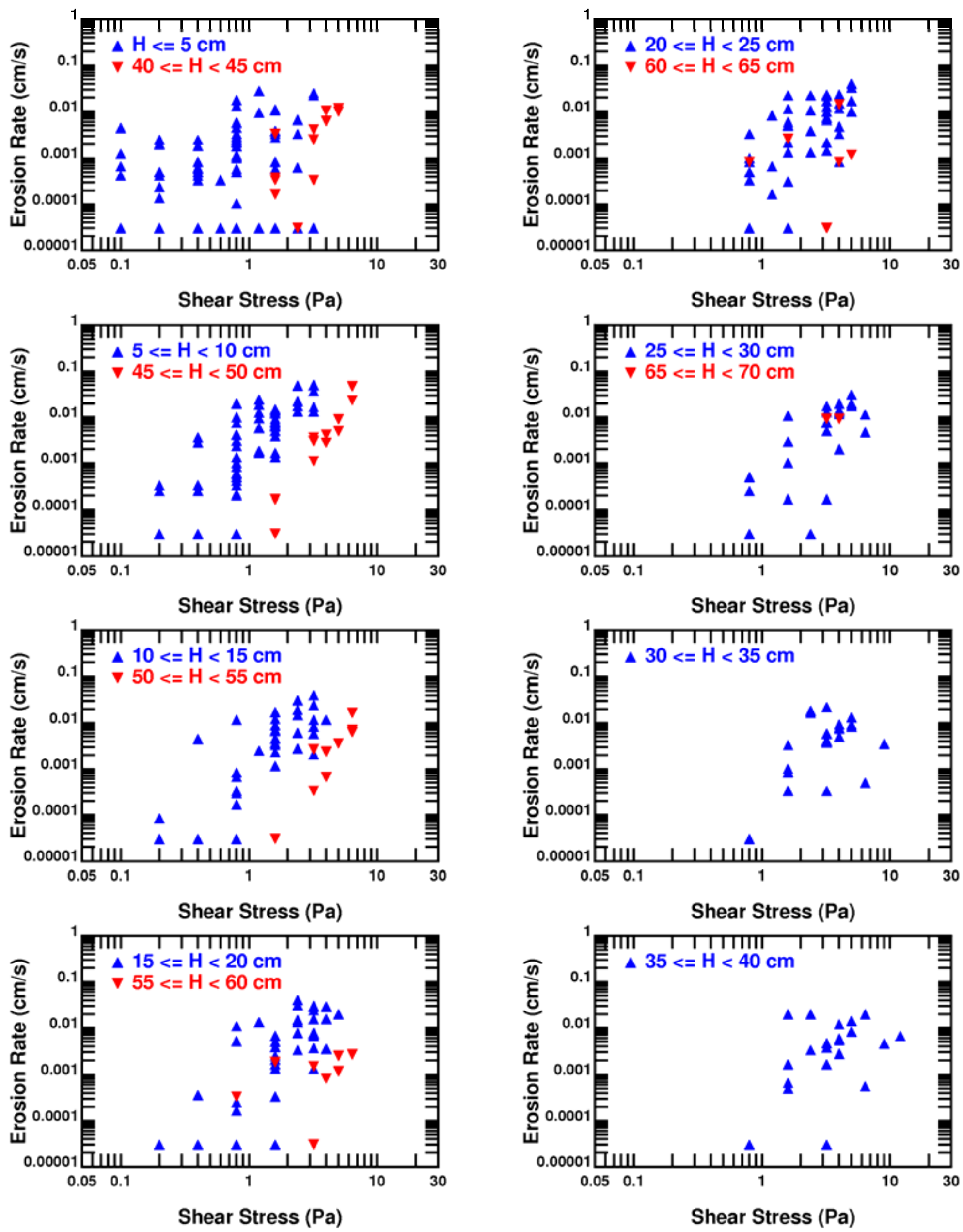


Figure GI-7-3b Erosion Rate Versus Shear Stress (as a function of depth in core) for Sedflume Experiments on Sediment with Median Particle Diameter (D_{50}) Greater than or Equal to 30 µm and Less than 100 µm

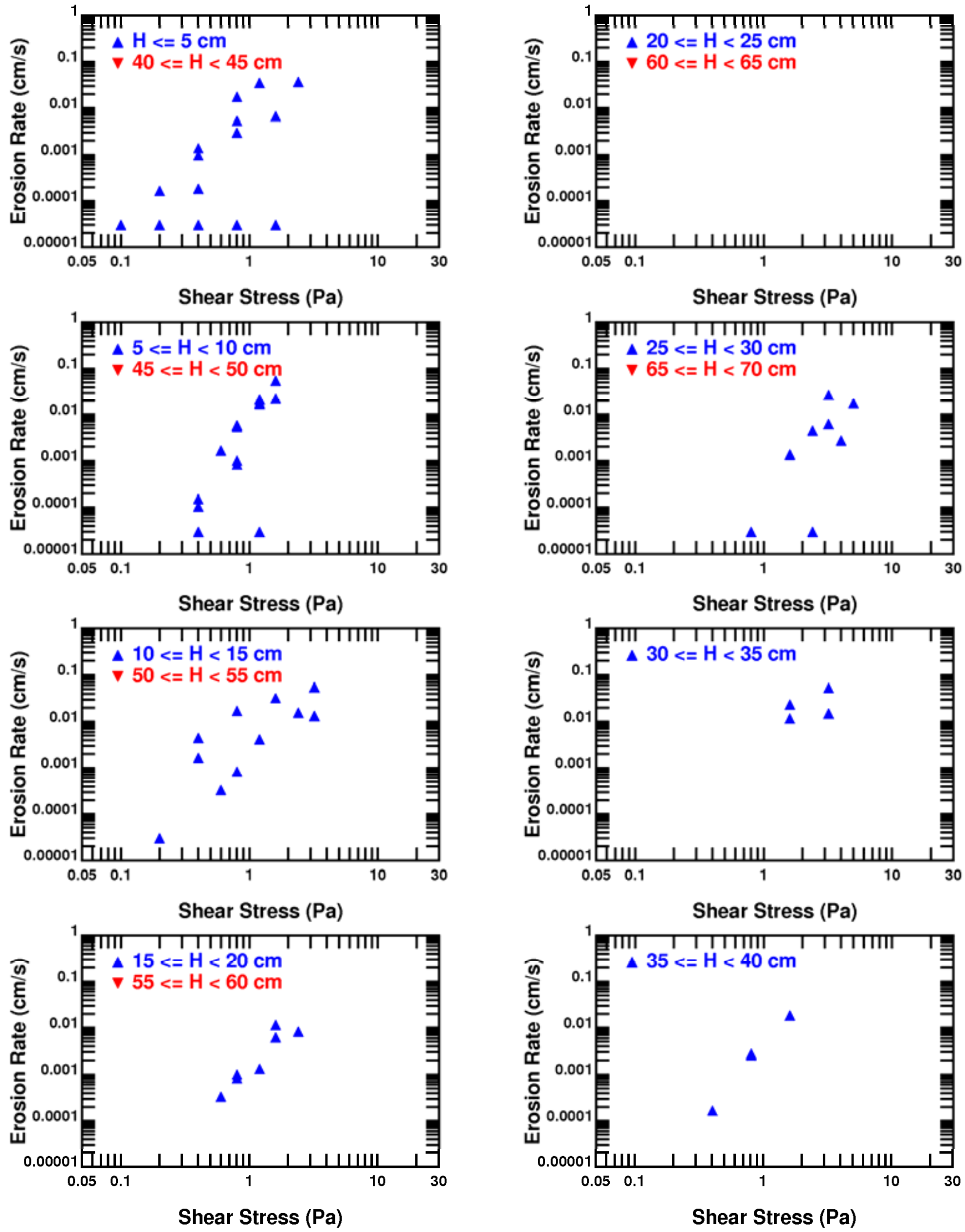


Figure GI-7-3c Erosion Rate Versus Shear Stress (as a function of depth in core) for Sedflume Experiments on Sediment with Median Particle Diameter (D_{50}) Greater than or Equal to 100 µm

Implementation of Dr. Lick's Simple Flocculation Model

In response to comments received from Reviewers, the flocculation model suggested by Dr. Lick was implemented in EFDC and tested. The implementation included an implicit iterative solver for the non-linear equation for floc diameter (Equation 4.4-36 of the Course Notes [Lick et al., 2005]). Floc diameters were then used to calculate settling velocities. Testing began with the simulation of a closed recirculating flume configuration (to mimic the flocculator used by Dr. Lick in his experiments), where the outflow and concentration at the exit section were fed back into the entrance section. This configuration results in steady-flow conditions.

The first series of tests started with constant suspended solids concentration and deposition and resuspension disabled, such that the suspension concentration remained constant. Predicted floc diameters for different concentrations and flow rates agreed with graphical material in Dr. Lick's lecture notes (Lick et al., 2005). For verification purposes, a stand-alone integration of the differential equation for floc diameter was performed, and the calculated floc diameters agreed with those predicted by EFDC when the residence time of the cells comprising the flume was relatively short, i.e., approximately 100 seconds (the residence time for cells in the Housatonic River is also on this order or less). For longer residence times (approximately 10,000 seconds), variability of the computed floc diameters from cell to cell was observed.

The second series of tests involved gradually increasing the suspended solids concentration by adding solids to the inflow, with the flow rate remaining constant. This emulated the addition of suspended solids to a recirculating flume. For slow rates of increase, on the order of 10 mg/L per day to 100 mg/L per day, the floc diameter changed gradually from that associated with the initial suspended solids concentration to that associated with the final concentration for the short cell residence time case. The same was true for the case when the concentration was reduced while the flow rate was held constant.

The third series of tests used the same recirculating flume configuration with resuspension and deposition activated. The test started with a small initial suspended solids concentration and a constant flow rate, which produced gradual resuspension. The flow rate was then increased by a factor of 10 over a time period of 1 day and then reduced to the original flow rate after another day to mimic a high-flow event hydrograph. The flocculation calculation predicted floc diameters and settling velocities as the flow rate increased and suspended solids concentration increased due to resuspension up to a point at which the flocculation calculation suddenly produced a settling velocity that resulted in almost instantaneous settling of suspended solids, producing a zero floc diameter that did not allow the floc diameter equation to continue to be integrated forward in time without some arbitrary re-initialization.

The fourth series of tests was conducted with the Housatonic River calibration configuration. Similar problems as in the third series of tests were encountered during moderate flow increases.

In summary, the flocculation model performed well under idealized conditions, i.e., without deposition and resuspension, when the residence time of the model cells is relatively short. This suggests that the flocculation model is calculating a sequence of steady-state conditions consistent with the constant concentration assumption inherent in the model formulation. However, the flocculation model failed under conditions of unsteady flow with resuspension and rapidly increasing suspended solids concentration, conditions that are obviously inconsistent with the constant concentration assumption. Although the flocculation model appears to have promise, it would require additional development to be sufficiently robust for application to the Housatonic River.

Evaluation of Sediment Transport Results Over Longer Time Periods

Several Reviewers recommended that sediment transport results from longer-term simulations be compared to additional types of data, such as estimates of sedimentation rates derived from vertical profiles of radioisotopes, and erosion and deposition patterns described by surveying of river cross-sections conducted multiple times since October 2000. EPA agrees that these are important comparisons and, as discussed in the Calibration Report, they will be included as part of the evaluation of the longer-term modeling, which will be presented in the Validation Report.

References:

Garcia, M., and G. Parker. 1991. Entrainment of bed sediment into suspension. *J. Hydraul. Eng*. 117:414-435.

Lick, W., J. Gailani, C. Jones, E. Hayter, L. Burkhard, and J. McNeil, 2005. The Transport of Sediments and Contaminants in Surface Waters. Short Course Notes, University of California, Santa Barbara. January 2005.

van Rijn, L.C. 1984. Sediment transport, Part II. Suspended load transport. *J. Hydraul. Eng.* 110(11):1613:1638.

Wu, W., S. Wanh, and Y. Jia. 2000. Nonuniform sediment transport in alluvial rivers. *J. Hydraulic Research* 38(6):427-434.

8. DIFFUSIVE FLUX

SUMMARY OF ISSUE:

Three Reviewers expressed concerns regarding the rate used in the model for diffusive flux of contaminants from the sediment to the overlying water, and the related topics of bioturbation and the depth of the active layer in the bed sediment (see General Issue 6). These Reviewers noted that the value used for this parameter was very high in comparison to values found in the literature. One Reviewer suggested that the value of this parameter should not be constant, but should vary with flow. Another Reviewer recommended a supplemental study to provide information that could be used to assign a value to this parameter.

RESPONSE:

In the Conceptual Site Model discussion (Section 4.5.2 of the final MFD), the processes of advection and diffusion are discussed. Both processes are included together in the concept of “vertical flux,” an analysis of this flux in the PSA is presented, and the lumped processes are concluded to be “. . . an important factor [for the] spatial distribution of water column PCBs . . .” This “vertical flux” is represented with a mass transfer coefficient, K_f, which can be thought of as a lumped parameter that accounts for several processes, in addition to diffusion. The flux from bio-irrigation, bioturbation, advection, and diffusion are subsumed into the sediment-water diffusive transfer term, which was calibrated to water column PCB data.

The value of K_f (1.5 centimeters per day [cm/d]), calibrated to Housatonic River data, is at the low end of the range of values summarized by Thibodeaux and others (Thibodeaux and Bierman, 2003; Thibodeaux et al., 2002). These authors noted intra-annual variations in K_f values in the Hudson, Grasse, and Fox Rivers, with lower values (3 to 10 cm/d) in the winter and higher values (20 to 40 cm/d) in the summer. One possible explanation for the Reviewers’ belief that the 1.5 cm/d rate used for the Housatonic River PSA is high relative to values used in other systems might be a summary presented in Thibodeaux et al. (2002), a reference cited by one of the Reviewers. In that paper, K_f values were presented based on PCB concentrations on particles. The authors tabulated K_f values in that way to facilitate comparisons with particle resuspension velocities. They divided the original K_f values by a factor ranging between 4150 and 8640 (the product of the partition coefficient and particle concentration), which resulted in values that are not directly comparable to the value of 1.5 cm/d used for the Housatonic River. When expressed in a way that is consistent with the manner in which the coefficient is used in the model (see Table GI-8-1), the magnitudes of the K_f values from these other sites are in the tens of cm/d – values comparable to, or in some cases considerably higher than, the 1.5 cm/d developed from calibration to Housatonic River data.

It is noted that the minimum value of 0.2 cm/d for the Fox River above DePere Dam comes from a model (Velleux and Endicott, 1994) that includes “background” resuspension at low-flow conditions, which the authors acknowledge “substantially influences water column PCB concentrations, but has little impact on solids concentrations.” Although there is no evidence that such an approach is more realistic or preferable for explaining the spatial pattern of increasing water column PCB concentrations than the “lumped parameter” approach, use of a “background resuspension” approach in the Housatonic River

Table GI-8-1

Summary of Mass Transfer Coefficients (after Thibodeaux et al., 2002)

Study	K_f (cm/d) (based on PCB concentration on particles)*		Product of Partition Coefficient and Particle Concentration	Original K_f (cm/d) (based on pore water concentration)	
	Minimum	Maximum		Minimum	Maximum
Fox River					
Above DePere Dam	2.60E-05	5.10E-03	7680	0.2	39.2
Below DePere Dam	5.20E-04	3.90E-03	7680	4.0	30.0
Hudson River					
Study 1	7.20E-04	9.60E-03	4150	3.0	39.8
Study 2	2.50E-03	6.00E-03	4150	10.4	24.9
Grasse	1.20E-04	8.10E-04	8640	1.0	7.0

* Values in mm/d in Table III of Thibodeaux et al. 2002, converted to cm/d.

modeling study would not be possible with the more mechanistic sediment transport approach in EFDC without adding multiple cohesive solids classes, for which data are not available.

An evaluation of the Housatonic River data was performed to determine if there were indications of intra-annual variations in the mass transfer coefficient (see pages B.4-32 through B.4-35 and Figure B.4-30 of the Model Calibration Report); however, no such patterns were observed. Therefore, the value of 1.5 cm/d was used for all months of the year and all flow conditions. One Reviewer suggested that K_f should be expected to vary with river flow based on changes in the thickness of the water column boundary layer. If the sediment-water PCB flux were strictly a diffusional process, this would be expected; however, the surface mass transfer coefficient, K_f, represents the flux from a number of processes, including bio-irrigation, bioturbation, advection, and diffusion. The spatial patterns of water column PCB concentrations reflect the combined effect of all of these processes, and do not suggest a variation with river flow.

The calibrated value of 1.5 cm/day produces agreement between simulated PCB concentrations and those measured at extremely low-flow conditions between 16 and 25 cfs (Figure B.4-37) and at above-average flow conditions between 109 and 244 cubic feet per second (cfs) (Figure B.4-38). As river flow and velocity continue to increase, the effect of non-particulate PCB transport on water column concentrations decreases because of increasing dilution, and particulate transport becomes increasingly important. Because the mass transfer coefficient, K_f, was calibrated to data collected at lower flow conditions when this transport process is more important, EPA believes that additional field measurements are not necessary to refine or further constrain this parameter.

Two Reviewers expressed concern that the calibration of the sediment-water column mass transfer coefficient, K_f, was based on reproducing the measured increase in water column PCB concentrations between sampling stations. In one case the concern was related to the spatial resolution of the segmentation of the river channel, which led the Reviewer to question if the PCB flux from the sediment that was represented in EPA's model by a "non-particle" flux might have been due to resuspension of bed sediment in a portion of the river cross-section. The Reviewer questioned if concurrent resuspension and deposition of sediment could be occurring in different parts of the cross-section, which would not be simulated in the model because the channel cross-section is, in general, represented by one grid cell. The model results reproduced the spatial profile of water column PCBs described by data collected during the lowest-flow surveys, when peak flows at Coltsville were between 16 and 25 cfs, and resuspension is neither expected, because of the low velocities, nor indicated by the spatial profile of suspended solids. Hourly flows at Coltsville exceed 25 cfs over 85% of the time. It is unreasonable to assume that substantial PCB accumulation could occur in locations that erode over 85% of the time; therefore, EPA believes that the "non-particle" flux represented by the mass transfer coefficient, K_f, equal to 1.5 cm/d, is a more reasonable explanation for the measured water column PCB profiles.

The second Reviewer who expressed concern over the calibration of the sediment-water column mass transfer coefficient, K_f, questioned whether EPA had adequately considered additional point and non-point sources of PCBs between the sampling locations. The only potential point source of PCBs in the PSA is the Pittsfield wastewater treatment plant (WWTP); however, sampling of sludge and effluent conducted by the City was non-detect for PCBs, indicating that the plant discharge is not a source of PCBs to the river. Those results are consistent with EPA's water quality sampling, which found no measurable increase in sediment or water column PCB concentrations downstream of the WWTP discharge. There is another mechanism in place under the direction of the Massachusetts Department of Environmental Protection (MDEP) to evaluate potential sources of PCBs in the watershed. This program did not identify any potential sources of PCBs located in the watershed (other than those that are known and are being addressed, such as Dorothy Amos Park on the West Branch) that would be expected to contribute runoff containing PCBs to the PSA.

References:

Thibodeaux, L.J. and V.J. Bierman. 2003. The bioturbation-driven chemical release process. *Environmental Science & Technology*, July 1, 2003. pp 252-258.

Thibodeaux, L.J., D.D. Reible, and K.T. Valsaraj. 2002. Non-Particle Resuspension Chemical Transport from Stream Beds. In *Chemicals in the Environment: Fate, Impacts, and Remediation;* ACS Symposium Series 806, American Chemical Society, Washington, DC., 2002. pp 130-149.

Velleux, M. and D. Endicott. 1994. Development of a mass balance model for estimating PCB export from the Lower Fox River to Green Bay. *J. Great Lakes Research* 20(2):416-434.

9. BANK EROSION

SUMMARY OF ISSUE:

Five of the seven Reviewers recommended that riverbank erosion and mass slumping be included as a process in the sediment/contaminant transport modeling. This process was variously termed “important” or “significant” by Reviewers.

RESPONSE:

EPA agrees that bank erosion/slumping is a process that should be included in the model, and, as discussed in the Model Calibration Report, has incorporated this process in the longer-term modeling (Phase 2 Calibration and Validation). EPA conducted several studies to collect the data that were used for estimating long-term rates of soil loss from the riverbanks.

Two different methods were used to develop estimates of soil erosion from the riverbanks in the PSA (see final MFD, pp. 7-8 to 7-11). One method provides long-term estimates and the other provides short-term estimates of soil erosion. The long-term erosion estimates were developed by overlaying georeferenced representations of the top of bank from aerial photographs of the river taken in 1952, 1970, 1990, and 2000. Changes in riverbank locations were quantified and summarized as average migration rates of riverbanks per year for the period 1952 through 2000. The mass of soil erosion was estimated from the average migration rate, height and length of eroding bank, and soil dry density. This analysis provides site-specific erosion estimates for each of the 69 erosion areas.

Estimates of short-term erosion rates were developed from measurements made in Reaches 5A (eight locations) and 5B (seven locations), which were selected to represent a range of erosion rates based on factors such as radius of the curvature of bends, bank vegetation, and soil characteristics. No areas of active erosion were observed in Reaches 5C or 6; therefore, no locations for short-term erosion rate estimates were selected downstream of Reach 5B. Three-dimensional terrain models were developed for each location from surveys conducted in November 2001 and June 2002. Differences between the two terrain models were used to estimate the average soil loss per linear foot of riverbank in Reaches 5A and 5B. These estimates were applied to the total length of eroding riverbank in each reach, which was determined from field observations.

These long- and short-term erosion estimates will be used to constrain the parameterization of bank erosion in the Phase 2 Calibration modeling. The effect of the representation of bank erosion/slumping on simulated suspended solids and PCB concentrations during the monitored storm events will be assessed through model-data comparisons.

One Reviewer suggested the use of a one-dimensional model for meandering streams to assess how much sediment enters the river through stream erosion. As of late August 2005, this model was still under development and was not available to the public. The approach adopted by EPA relies on the short- and long-term erosion estimates based on site-specific information. If a model were used to supplement EFDC by generating soil erosion estimates, those estimates would be evaluated by comparison with the short- and long-term erosion estimates. Because the erosion estimates that will ultimately be used will be consistent with the site data, EPA believes it is reasonable to base the model inputs on the site data, rather than a supplemental one-dimensional model.

EFDC does not have the capability to calculate planform evolution and EPA does not plan to account for this process, other than through the representation of bank erosion/slumping. The analysis of aerial photographs of the river indicates that the river has not changed course substantially over the last half-century. Representing the erosion/slumping of bank soil and associated PCBs captures the primary impact of bank erosion.

10. VARIABILITY IN MODEL RESULTS VS. DATA

SUMMARY OF ISSUE:

Five Reviewers questioned that the variability in the model results was less than the variability in the data, but the Reviewers differed in their conclusions regarding how this decrease in variability would affect the utility of the model for evaluating remedial alternatives. One Reviewer indicated that the model should not be expected to reproduce the variability in the data and that this would be a potential problem only for processes that vary non-linearly with contaminant concentration. Another Reviewer considered the decrease in variability (referred to as a “filtering” of the data) to be a form of bias, but did not indicate it to be a major problem. A third Reviewer, however, considered the inability of the model to generate results with variability equivalent to that of the data to be a “serious problem.”

RESPONSE:

In their comments, Reviewers used the terms “variance” and “variability” interchangeably. Although these concepts are related, they are not identical. The term variance, whether used in the strict statistical sense referring to the mean squared deviation from the mean of individual measurements, or in a more general statistical sense, is a measure of the variability (spread) of the data, and incorporates the concept of stochasticity or uncertainty. In the context of a modeling study, the variance of a parameter in a specific grid cell and time step would be represented by a distribution rather than a single deterministic value. The distribution would be characterized by a measure of central tendency (e.g., mean) and a measure of dispersion (e.g., variance), along with other descriptors such as skewness and kurtosis.

Variability, on the other hand, refers to the range of values for an individual parameter in a specified time and space, even when each individual measurement or value of the parameter is made or calculated without any uncertainty (i.e., is deterministic). Translated into modeling terms, variability is the range of values of a parameter over a number of model grid cells and/or a number of time steps. Each grid cell/time step has only a single deterministic value for the parameter, but taken together the deterministic individual values have an associated range and variance. The remainder of this response addresses the concepts of variance and variability separately.

Regardless of whether the concept of variance or variability is the issue, EPA disagrees that a difference (reduction) in variability of the model results compared to the data represents a form of bias, and believes that the Reviewer was not using the term "bias" in the traditional sense. The distinction is important, however, because consistent under- or overprediction by the model (bias) could affect the response time of PCB concentrations in the PSA to remedial alternatives and complicate the comparison of various alternatives to each other and to natural recovery. Reviewers were in agreement that there is no indication of this type of bias.

Variance

In order for a model to truly provide a measure of variance in its output, the model must be probabilistic rather than deterministic. A probabilistic model, when run for multiple trials, does not always return the same value of an output parameter for a specific set of inputs. Instead, numerous runs of the model produce a distribution of results, which can be used to make certain inferences regarding the ability of the model to produce results that duplicate the system being modeled. Such distributions can also potentially be used to conduct statistical hypothesis testing of model results.

All three of the models comprising the Housatonic River modeling framework are deterministic, as are models of sediment/contaminant fate and transport generally. That means that, for a given set of model process formulations, state variables, and input parameters, the model will always return the same value for each parameter at each time step and grid cell (i.e., there is no distribution or spread of outputs for a given place and time). The only way to produce such a distribution is to run the models in a probabilistic mode, supplying distributions rather than fixed values for the input parameters, and sampling from the distributions in the general manner of a Monte Carlo analysis. In principle, it is possible to do that provided realistic input distributions can be developed for the parameters of interest (which is not necessarily simple, or even possible). However, the computational demands of EFDC preclude such a probabilistic application. For HSPF and FCM, an uncertainty analysis will be conducted using probabilistic methods, but the uncertainty analysis for EFDC will require a different approach (see the response to General Issue 11 for a discussion of the uncertainty analyses).

In the absence of an explicit determination of the distribution, or variance, associated with each parameter for each model cell at each time step, it is possible to approximate the variance by assuming that the variability of results will be approximately equal to the variability in the input data when interpreting the model output. However, the same variance may not be applicable for longer-term runs in which the central tendency, or simulated, values for many parameters will show a greater difference at the end of the run than at the beginning. In that eventuality, scaling of the variance to the mean may be more appropriate

Variability

Although a number of Reviewers commented on the variability (and/or variance) of model output in their comments on the Model Calibration Report, similar concerns were not expressed in the comments received on the MFD document. In fact, only one Reviewer even mentioned variability of model output, stating:

> "Models of PCB transport and fate, and underlying theory, are not sufficiently robust that parameter values determined a priori can account for all of the site-specific variability that is observed in critical model parameters. This is not a weakness of the models specifically, rather an acknowledgment that all transport and fate models are imperfect representations of chemical behavior in an extremely complex system." (D. Endicott, May 24, 2001)

EPA agrees with the Reviewer's assessment and recognizes that the output of the Housatonic River model does not reproduce the range of variability in the data used to develop the model input. Rather than a fault of the modeling framework or the models themselves, such a result is an inevitable and expected consequence of any modeling activity. Models are only representations of reality, and are not intended to include all of the components of the system being modeled, or of all the relationships between them. Indeed, if the model and the system being modeled are identical in all of their characteristics, the concept of modeling loses its usefulness (Gold, 1977). Therefore, it follows that models must be a simplification of the system being modeled, and simplification necessarily results in the inability of any model to completely simulate either the inputs to, or the outputs from, that system.

The range of values (the variability) of any parameter in the environment is a function of the range of values for the other parameters that affect the parameter in question. For example, surface runoff is a function of precipitation and land use (among other variables). In reality, any type of land use is a continuum that does not neatly fall into a discrete number of categories that in turn have two conditions, permeable and impermeable. In order to develop a model of surface runoff successfully, the continuum of land use must be simplified into a number of categories, and it is not possible for the categories to capture the extremes of the range of permeabilities. Accordingly, the range of surface runoff values calculated by the model will necessarily be truncated and will not span the range of variability in the original data used to parameterize the model. Increasing the number of input parameters (most models use substantially more parameters than this simple example), will in most cases increase the amount of attenuation of variability.

The question of importance for the modeling study is not whether the results of the model span the full range of variability in the input data, but whether certain features of the data that are important considering the goals of the modeling study are reproduced. Central tendency, or mean, predicted values averaged over a specified temporal and spatial scale are one obvious feature that should be reproduced with reasonable accuracy, as is the general pattern of variability. In the case of the Housatonic River modeling study, it is important for some parameters (e.g., flow) that the magnitude and frequency of peaks (but not necessarily the timing) be reproduced with some reasonable degree of fidelity; however, for other parameters (e.g., PCB concentrations in sediment) simulating the full range of variability is of relatively little importance because receptors integrate their exposure to various contaminant concentrations, and this is what is of interest to satisfy the goals of the modeling study.

EPA agrees with the comment that the model cannot be expected to predict the variability observed in the data and that the use of mean concentrations does not pose a problem because sediment-water exchange of PCBs is represented in the model by dissolved and particulate exchange that varies linearly with concentration.

Reference:

Gold, H. 1977. *Mathematical Modeling of Biological Systems*. J. Wiley & Sons, New York, NY. 357 pp.

11. ANALYSIS AND EVALUATION OF MODEL UNCERTAINTY

SUMMARY OF ISSUE:

Six of the seven Reviewers provided comments on the approaches and methods used to quantify and evaluate model uncertainty. The Reviewers correctly noted that uncertainty analyses have been conducted only for FCM at this time, with analysis of uncertainty for HSPF and EFDC and for the linked modeling framework deferred until the Validation Report. Some of the Reviewers questioned whether it would be possible to conduct an uncertainty analysis for EFDC, primarily because of limitations arising from computational requirements. Several Reviewers noted that uncertainty arises from different sources (e.g., variability in data, uncertainty in model formulations).

RESPONSE:

As indicated by the Reviewers, only the uncertainty analysis for FCM was included in the Model Calibration Report. The uncertainty analyses for HSPF and EFDC, and the linked uncertainty analysis to investigate the propagation of uncertainty through the three models, were deferred and will be included in the Validation Report. There are no generally accepted methods for conducting uncertainty analyses for numerical transport and fate models, and EPA was looking for input and suggestions from the Reviewers prior to expending the considerable resources necessary to conduct uncertainty analyses involving HSPF and EFDC. Because the bioaccumulation model is comparatively simple and inexpensive to run, it was considered feasible to include a preliminary uncertainty analysis for FCM in the Model Calibration Report. This overall approach to analysis of uncertainty and the underlying reasons for adopting it were presented to the Peer Reviewers by EPA at the document overview meeting on April 13, 2005.

Uncertainty Analysis for Watershed Model (HSPF)

The details of the proposed approach to conducting the uncertainty analysis for HSPF were presented in Section A.6.2 of the Model Calibration Report. Reviewers did not comment directly on this proposed approach, but did provide comments relative to conducting uncertainty analysis that were generally consistent with the proposed approach for HSPF. Accordingly, the uncertainty analysis will be conducted substantially as described in the Model Calibration Report, with some refinements that will facilitate the analysis of uncertainty as it propagates through the linked models.

The uncertainty analysis for HSPF will be conducted for the Phase 2 Calibration period (1990 through 2000) using a Monte Carlo approach involving approximately 500 iterative model runs, with selected model parameters being chosen from assigned probability distributions. The model results will then be processed for the same output variables and locations used for the sensitivity analysis presented in the Model Calibration Report (Section A.6.1), and presented as measures of central tendency and confidence intervals for each output variable and location. The parameters to be varied for the HSPF Monte Carlo analysis will be the same 28 parameters evaluated in the sensitivity analysis.

Uncertainty Analysis for Hydrodynamic-Sediment/Contaminant Transport Model (EFDC)

The plan for uncertainty analysis of EFDC was presented in Section B.5.3 of the Model Calibration Report. Based on comments received from Reviewers, and additional consideration by the modeling team of how the analysis of uncertainty could be conducted for a complex and computationally demanding model such as EFDC, EPA has further developed the approach to the EFDC uncertainty analysis beyond that outlined in the Model Calibration Report. Because it is computationally less intensive and conceptually simpler to understand, the Kolmogorov-Smirnov (K-S) confidence limits approach will be used to evaluate EFDC uncertainty. This approach is discussed in detail in Attachment 2 (Kolmogorov-Smirnov Approach to Evaluation of Uncertainty for EFDC).

EPA agrees with the Reviewer who recommended not conducting the “bounding” analysis proposed by GE in their presentation at the Peer Review public meeting. Although such an approach has appeal due to ease of calculation and presentation, EPA believes that the proposed analysis has a number of fundamental problems. Chief among these are issues related to the selection of parameters to use in the analysis, difficulties associated with specifying true ranges for the parameters selected, and the problem of non-orthogonality. In combination, these and other issues lead to results that are not fully objective and do not represent true uncertainty bounds (in a statistical sense) on model predictions.

Uncertainty Analysis for Food Chain Model (FCM)

The procedure for conducting the uncertainty analysis for FCM will be identical to that presented in the Model Calibration Report. In addition, the FCM uncertainty analysis will incorporate the propagated uncertainty of the parameters output from EFDC, as described below.

Propagation of Uncertainty through Linked Models

The procedures outlined above provide a framework for conducting an uncertainty analysis that will evaluate the propagation of uncertainty through the linked models. Analysis of uncertainty for the linked models will focus on those parameters that are output from one model and used as input to another model: flow at the upstream boundary of the PSA (HSPF to EFDC) and PCB concentrations in sediment and on particulate organic matter (POM), respectively (EFDC to FCM). The uncertainty in the output for each of these parameters, as quantified by the approaches discussed above, will be used to describe the error distribution about each parameter when it is input to the next model in sequence.

Each of the 500 HSPF Monte Carlo simulations will produce a time-series of hourly flows for the 1990 through 2000 Phase 2 Calibration period. This group of model results will be ranked based on the effect of each time series on the EFDC model simulation – i.e., average flow over the time period, or some similar parameter, will be evaluated for each time series and used to create a distribution. The 5^{th}-percentile and 95^{th}-percentile time series from this distribution will be used in the K-S analysis to provide bounding estimates on the EFDC results. These EFDC bounding estimates can then be used to parameterize a distribution for each of the EFDC output parameters that are passed to FCM, thereby propagating uncertainty through the series of three linked models.

12. DEMONSTRATIONS OF MODEL PERFORMANCE

SUMMARY OF ISSUE:

All of the Reviewers made some comments that suggested ways that the models could be improved or tested to better demonstrate model calibration and/or correct representations of processes. Examples of these (which are not represented by another General Issue) include:

- Plots of measured versus simulated TSS as an “equivalence” plot.
- Conduct a “trial” remediation (although this is not possible in a real sense, it can be simulated) and evaluation of the model response.
- Comparison of velocity data to simulated velocities.
- Compare simulated erosion and deposition patterns to the data showing locations and changes in bed elevation (including Woods Pond).
- Show the proportion of PCB burden in different species/trophic levels attributable to water-related versus sediment-related sources.
- Evaluate PCB fluxes between the river and floodplain during overbank flow conditions.
- Use statistical methods such as mean squared error or confidence limits in making model/data comparisons to supplement the uncertainty analyses.

RESPONSE:

- EPA agrees that plots of measured vs. simulated TSS would be useful and will include these figures in the Model Validation Report.
- EPA agrees that simulation of an example remediation would be useful as a demonstration of model performance and will consider inclusion of one or more such simulations in the Model Validation Report.

- Figure GI-12-1 compares the simulated versus measured velocities. The differences between the simulated results and the data are generally within 15 to 20% at Holmes Road. The model results are in very good agreement with the data at New Lenox Road, and somewhat lower than the data at Woods Pond Footbridge.

- With regard to the latter, the difference tends to increase with increasing velocity. As noted in the discussion of Figure B.2-37 in Section B.2.3.2.5 of the Model Calibration Report, the discrepancies between the simulated and measured velocities at Woods Pond Footbridge are attributed to size of channel passing through the bridge abutment relative to the scale of a model grid cell.

- EPA agrees that the comparison of simulated erosion and deposition patterns to the data on changes in bed elevation would be valuable, but believes that it is more appropriate to conduct such an evaluation using the results of the long-term (Phase 2 Calibration) period. The results of the evaluation will be presented in the Model Validation Report.

- EPA agrees that an evaluation of the sources of PCB burden in different species/trophic levels will provide insight into the relative importance of water- versus sediment-related sources. This analysis will be performed and presented in the Model Validation Report.

- The simulated net flux of PCBs to the floodplain during periods of out-of-bank flows is shown in Table GI-12-1.

Table GI-12-1

PCB Flux

Reach	(kg/year)
5A	2.8
5B	4.1
5C	4.8
6	0.5

- The approach to conducting the uncertainty analysis for the modeling study has been revised and is discussed in detail in the response to General Issue 11 and in Attachment 2 to this responsiveness summary. The revised approach includes the type of statistical analyses recommended by the Reviewers.

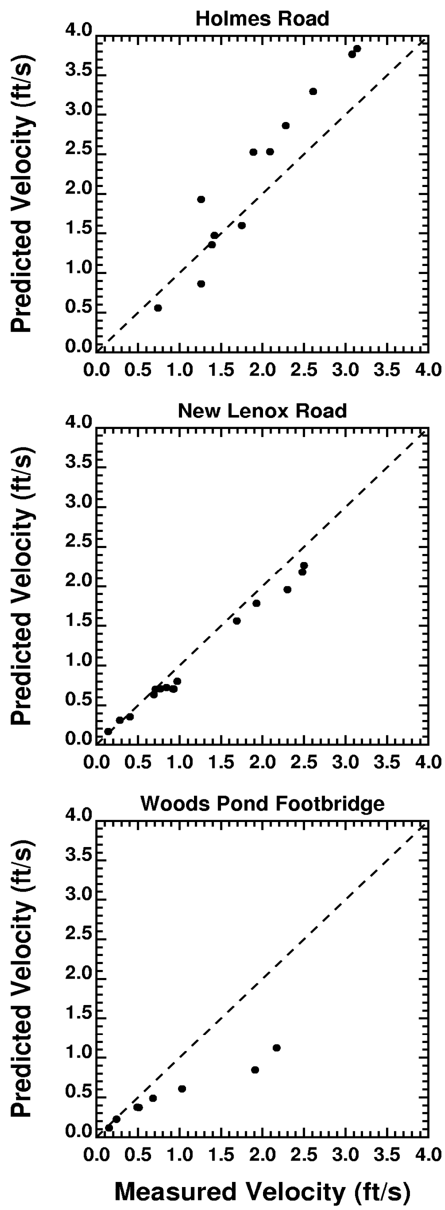


Figure GI-12-1 Comparisons of Simulated vs. Measured Velocities at Holmes Road, New Lenox Road, and Woods Pond Footbridge

RESPONSE TO SPECIFIC COMMENTS

OVERVIEW COMMENTS

W. Frank Bohlen:

Introduction

Agreements developed between the General Electric Company and the U.S Environmental Protection Agency regarding means to optimize remedial activities intended to reduce or eliminate environmental and/or public health effects thought to be associated with exposure to PCB contaminated sediments include the development of a predictive numerical fate and transport model for application in the Housatonic River system. Given the complex of processes affecting this transport, this represents an ambitious project. The proposed model consists of three primary components, a watershed model (HSPF), a hydrodynamic/sediment-contaminant transport model (EFDC) and a bioaccumulation model (FCM). The models are linked but non-interactive. Details of each of these models and the results of the calibration phase are presented in a lengthy master volume and three detailed Appendices. A Peer Review Panel was asked to review this work and to answer a series of specific questions. Before I get to these questions I'll begin with a number of more general observations and recommendations.

General Comments

As discussed by several reviewers, assessments of model adequacy require a clear understanding of the intended application. Beyond the fact that the developing framework is to be used in the assessment of remedial alternatives, little detailed information is provided. It is not clear whether the model is to be used to assess benefits of one scheme relative to another or to provide an absolute assessment of selected schemes. One might argue that such specification is premature and in fact will be based on model results detailing the relative importance of particular source areas or processes. While this may ultimately prove to be the case, it still would be well to begin with a defined set of possibilities. In the case of the Housatonic River Study Area (PSA) this seems entirely possible due to the depth and breadth of data available detailing all primary site characteristics and the associated contaminant distributions. We know, for example, that a significant fraction of the total PCB mass in the river downstream of the confluence to Woods Pond is located in the floodplain. Given our less than perfect understanding of the range of processes affecting contaminant transport to and from the floodplain this fact might suggest that models would be best used to assess relative benefits so as to favor cancellation of errors. Alternatively, one might argue that the floodplain is primarily a trap and it is the more mobile components of the system such as the stream bed sediment column that is responsible for continuing exposure and the downstream flux of PCBs. This system, while still complicated may be more amenable to absolute assessment. Future discussions of model adequacy would benefit from a clear concise definition of the most probable primary use of the suite of models.

RESPONSE O-FB-1:

Please refer to the response to General Issue 1.

Beyond care in the definition of model application, this entire exercise would benefit from a careful (i.e. brutal) editing of the reports in the interest of clarity, understanding and the retention of the reader's interest. This is a complicated multi-faceted effort with a large number of investigators spread across the country (at least). I recognize that the coordination of the resulting writing effort is itself something of a herculean task. This must be faced however, if the goal is to produce a product that is at once comprehensive and amenable to detailed review and evaluation.

RESPONSE O-FB-2:

The U.S. Environmental Protection Agency (EPA) recognizes that the Model Calibration Report (WESTON, 2004), particularly the technical appendices, presents a large amount of highly technical information and that it may be difficult for some readers to process. In fact, the document did undergo careful and iterative technical editing to improve consistency, clarity, and readability. However, EPA also needs to ensure that the scientific analyses performed by EPA are thoroughly documented.

EPA will direct more effort toward streamlining and formatting the Model Validation Report to improve readability.

Reference:

WESTON (Weston Solutions, Inc.). 2004. *Model Calibration: Modeling Study of PCB Contamination in the Housatonic River*. Prepared for U.S. Army Corps of Engineers and U.S. Environmental Protection Agency.

The Executive Summary provides the framework to build on. It would however benefit from additional detail regarding boundary conditions applied for each model. These were spread through each of the individual model presentations and as such, easy to loose track of. A summary statement up-front would help.

RESPONSE O-FB-3:

Decisions regarding the type and amount of information, and the extent of technical detail, presented in an Executive Summary are always difficult. EPA's intention was to prepare an Executive Summary that would be useful to managers and other non-technical readers, and it was believed that such readers would not require technical information regarding boundary conditions. The Reviewer's comments will be considered during the preparation of the Model Validation Report.

The individual sections in Volume 1 dealing with each model would benefit from editing by a single hand to provide a consistent "story line" throughout. In general I found that the Appendices added little to the discussion. Much of the material presented was already available in Volume 1. Future editions might use the Appendices as the site for the majority of the data plots and in some few cases for detailed elaboration of some particular aspect of model formulation. This combination of efforts would reduce the overall size of the report and contribute to the acceptance of the model predictions by the broadest possible user group.

RESPONSE O-FB-4:

The Model Calibration Report technical appendices were necessarily prepared by the organizations responsible for each of the component models of the modeling framework. The appendices were extensively edited to improve readability and to present the material in a consistent manner. However, achieving consistency was not entirely possible because of the different structure of the individual models, which dictated somewhat different approaches to demonstrating calibration.

The Reviewer is correct that there is some repetition of information between Volume 1 and the technical appendices. This was intentional and was done with the understanding that few readers would likely read the Model Calibration Report in its entirety. It was assumed, however, that some individuals, such as the Peer Reviewers, would require far more detailed information than could be summarized in Volume 1. EPA will consider alternative approaches to formatting the Model Validation Report, including the Reviewer's suggestion.

Douglas Endicott:

As a reviewer, it is interesting to see how much this project has progressed in the 4 years since the review of the Framework document. The Housatonic River is a challenging system to understand and model. The PCB transport and fate processes in this river are unique in a number of ways, some of which I do not fully understand. As I mentioned in my preliminary comments, I am very impressed by the modeling work that is presented in the calibration report. I think that the project team has done a thorough job in assembling a suite of models that address the major processes affecting PCBs in the Housatonic River, and along the way have overcome a number of obstacles presented by site-specific data that challenge conventional wisdom. I remain optimistic that the modeling tools under development here will be valuable in terms of forecasting the outcome of remediation alternatives. I am also hopeful that concerns raised during the Calibration peer review will be considered carefully be EPA, and used to guide refinement of the modeling tools.

The modeling framework which we reviewed 4 years ago has changed considerably since then. The AQUATOX model biological/food chain component has been abandoned, and the QEA FOODCHAIN model has taken its place to predict PCB bioaccumulation. EFDC is now the sole framework for PCB transport and fate, although a number of potentially significant processes appear to be missing from this model. The role of HSPF in simulating the flow and solids boundary conditions to the PSA has been scaled back considerably. And perhaps most significantly, the modeling team has implemented EFDC using a 1-dimensional segmentation scheme in the water column for most of the PSA.

RESPONSE O-DE-1:

EPA agrees with the Reviewer's synopsis of the major features of the modeling framework and the changes that have been made as a result of the first Peer Review. Discussion of processes that may need to be modified or added to the modeling framework is provided in several other areas of this Responsiveness Summary, e.g., responses to General Issues 5, 6, 7, 8, and 9.

Calibration datasets have been generated for stage and flow, solids, organic carbon and PCBs in water, sediment, and biota compartments. An enormous number of sediment samples have been collected and analyzed to define initial conditions and characterize the variability of PCBs, organic carbon, grain size, porosity, etc. in the sediment bed. Unfortunately, much of the apparent variability in this data remains unexplained. Hints are provided in the RFI Report and Appendices that major components of this variability may be attributable to either measurement errors (including interlaboratory error) or the judgemental/focused bias applied in much of the sediment sampling. Variability arising from these factors was estimated, and could be used to better evaluate the sediment data.

RESPONSE O-DE-2:

There is no indication that either interlaboratory (measurement) error or focused selection of sampling locations is a major component of the observed variability in the PCB data, although each is recognized to be a potential source of bias. As referenced in the RFI Report (BBL and QEA, 2003), the field laboratory used by EPA for analysis of the majority of the samples from the intensive 1998 and 1999 field program was recognized to have lower extraction efficiency than the fixed laboratories used for earlier studies. This bias, estimated to be approximately 30%, admittedly complicates analysis of temporal trends, but is a minor source of variability.

Similarly, EPA recognized that focused sampling of known or suspected areas of higher contamination could potentially result in calculation of biased measures of central tendency. This is one reason that spatial averaging was used for calculation of exposure frequencies in the Human Health Risk Assessment (HHRA) for Rest of River. Again, such potential bias is only a minor contributor to variability.

The contribution of analytical variability to the observed total variability in the PCB data was evaluated in detail and reported in Attachment 4 to the Calibration Report. There is no indication that the magnitude or sources of such variability are unusual in any way for this type of study, and analytical variability is also a relatively minor contributor to the total variability in the data, resulting in the conclusion that the observed two to three order-of-magnitude small-scale variability in sediment PCB concentrations is a real feature of the river sediment. Although EPA agrees that the lack of an explanation for this variability is somewhat unsatisfying, it is not necessary to develop such an explanation to achieve the goals of the modeling study; therefore, additional evaluation of the sediment data to better explain this phenomenon is not necessary. Please refer also to the response to General Issue 5.

Reference:

BBL (Blasland, Bouck & Lee, Inc.) and QEA (Quantitative Environmental Analysis, LLC). 2003. *Housatonic River – Rest of River RCRA Facility Investigation Report.* Volume 1. Report. Prepared for General Electric Company, Pittsfield, MA.

The observation of extreme spatial variability in sediment PCB concentrations, yet no apparent temporal trends is perplexing.

RESPONSE O-DE-3:

Please refer to the response to General Issue 5.

Manipulations of the sediment data are in some cases (e.g., organic carbon content of sediment particle size fractions and TOC normalization of PCB concentrations) so complex and torturous that their descriptions are unintelligible.

RESPONSE O-DE-4:

EPA used complex analysis techniques, including maximum likelihood estimation methods designed for censored data and log-normal distribution theory, to generate the best estimates for model inputs based on site data. The detailed descriptions of these analyses were included to meet the objective of transparency. EPA acknowledges that, like the analyses, the descriptions are complex. Given the complexity, it is appreciated that considerable effort would be required for a reader to grasp the details of the procedures.

Additional analyses were also conducted in an effort to explore possible explanations for some of the variability observed in the sediment data, most of which were unsuccessful. These analyses were presented in the Model Calibration Report to address potential questions regarding what was done to try to understand components of the variability.

Although observations of erosion and deposition were made at a series of transects, this important information has not been directly utilized in the model calibration. Likewise, bank erosion was monitored but this data was not used in model calibration.

RESPONSE O-DE-5:

Please refer to the responses to General Issues 9 and 12.

A less extensive sampling and measurement program was carried out for the water column, primarily based on monthly and event sampling at 5 fixed locations. The resulting data provide a good sense of the limited spatial trends in water column PCB concentrations through the PSA at low and moderate flow rates. These data were used primarily to calibrate a diffusive flux from sediment pore water. Because of the limited duration of monitoring in the water column, relatively few large flood events were sampled. In this and some other important ways (e.g., the boundary conditions and water column PCB partitioning) water column monitoring was limited to the point that the calibration suffered. In at least one case (the suspended solids composition at the upstream boundary), the modeling team creatively overcame this limitation.

RESPONSE O-DE-6:

The objective of the storm event sampling program was to capture large flood events. The magnitudes of the flows monitored during this program reflect the difficulty in executing a storm event sampling program. EPA and its contractors made every effort to be creative in obtaining as much information as possible regarding large flood events. Sampling of large flood events was complicated by two factors: (1) limitations associated with weather forecasting that resulted in events for which the sampling crew was mobilized, but which ultimately did not occur as predicted; and (2) storms that occurred but were not predicted with enough lead time to mobilize the sampling crew. A sampling crew of 28 was necessary to perform all of the tasks in the storm event sampling standard operating procedure (SOP) over a 24-hour period for a 2- to 5-day event duration (sampling had to continue until at least some samples were obtained on the falling limb of the hydrograph; the larger the storm, the longer the duration). The second factor was the cost of conducting storm event sampling. Mobilization alone for a single event cost more than $30,000, and conducting a full storm event protocol cost more than $120,000. In some cases, as is typical with such a large sampling program, some deviations from the SOPs occurred that necessitated careful analysis and use of the data and/or other approaches to resolve gaps in the data record. Although it would have been desirable to have monitored a larger flood event, the decisionmaking for Rest of River could not be delayed to obtain data from a storm with flows that occur with 10- to 100-year return intervals.

Monthly sampling (which is predictable and far less resource-intensive) has been performed continuously since 1996 (more limited data are available from earlier sampling efforts) and continues at approximately 10 stations, to develop a long-term record of water column data for both the PSA and areas further downstream. These data are very useful for evaluating spatial and temporal patterns. This program has been continued, as well as the monitoring associated with remediation and biennial fish sampling, in part in response to the comment from this Reviewer on the Modeling Framework Design (MFD) document (WESTON, 2004).

In addition, three water column PCB partitioning data sets were collected, at low-, intermediate-, and high-flow conditions.

Reference:

WESTON (Weston Solutions, Inc.). 2004. *Modeling Framework Design: Modeling Study of PCB Contamination in the Housatonic River*. Prepared for U.S. Army Corps of Engineers and U.S. Environmental Protection Agency.

To support calibration of the food chain model, biota were sampled and analyzed for congener-specific PCBs. Target fish species were sampled in 4 of 5 reaches in the PSA. Fish in each reach were treated as discrete populations, based on habitat and life history assessments. The data provide good measures of the variability in PCB concentrations between reaches and age classes of fish. The absence of data for the 5th reach (5D) is unfortunate, because the highest PCB concentrations were predicted by the bioaccumulation model for fish in this reach. The lower food chain (trophic levels 1 and 2) were underrepresented in the biota sampling effort.

RESPONSE O-DE-7:

EPA agrees that the fish tissue data from the field collections provide good overall characterization of differences in PCB bioaccumulation among river reaches and fish age classes. For most species and reaches, sample sizes were sufficient to characterize the relationship between PCB concentration and fish age, and also to demonstrate the variability of individual fish concentrations within local subpopulations.

During the period of greatest sampling intensity (fall 1998 EPA sampling), the river conditions prohibited passage of the electro-shocking boat into the backwater areas. The resulting lack of Reach 5D fish tissue data increases the uncertainty of model calibration for this reach. However, as indicated by another Reviewer (see Comment 1-FG-20 below), one of the objectives of the model is to estimate bioaccumulation concentrations where concentration data are not available.

EPA agrees that lower food chain levels (i.e., invertebrates, periphyton) were not sampled as intensively as fish. However, given that the Housatonic River exhibits large variability in sediment PCB concentrations over small spatial scales, emphasis on fish species that integrate their exposures over larger areas of sediment was considered appropriate and a more reliable means of evaluating model performance, and also provided necessary data for the risk assessments. The relatively low sample sizes of invertebrate tissues in some individual sampling events were offset by the multiple lines of evidence available for assessing bioaccumulation in lower trophic species. For example, model simulations for sediment epifauna were compared against tissue PCB concentrations measured in D-net invertebrate samples, in crayfish samples, and in tree swallow stomach contents (i.e., representative of emergent aquatic insects). In addition, model-generated biota-to-sediment accumulation factors (BSAFs) were compared against literature values for freshwater invertebrates. These comparisons provided evidence that model simulations of lower trophic level organisms were reasonable.

Additional sampling and experimental activities were conducted to support various aspects of the models (SEDFLUME, bed load, pore and surface water partitioning). However, a number of other fairly standard water quality measurements were not conducted, including point and nonpoint source PCB monitoring. Supplemental studies to constrain several ambiguous parameters (e.g., pore water PCB diffusion flux and vertical extent of sediment mixing) appear to be necessary to support the calibration.

RESPONSE O-DE-8:

The Rest of River modeling study is supported by extensive and intensive data collection efforts, greater than what is typically performed at a hazardous waste site. These efforts included the collection and analysis of water column samples as part of routine water quality monitoring and under many different flow regimes, including during large storms with out-of-bank flows.

The only point source in the Primary Study Area (PSA) is the Pittsfield wastewater treatment plant (WWTP); analyses of sludge and effluent conducted by the City of Pittsfield were non-detect for PCBs, indicating that the plant discharge is not a source of PCBs to the river. Those results are consistent with EPA's water quality sampling, which found no measurable increase in sediment or water column PCB concentrations downstream of the WWTP discharge. In addition, there is another mechanism in place under the direction of the Massachusetts Department of Environmental Protection (MDEP) to evaluate potential sources of PCBs in the watershed. This program did not identify any potential sources of PCBs located in the watershed (other than those that are known and are being addressed, such as Dorothy Amos Park on the West Branch) that would be expected to contribute runoff containing PCBs to the PSA.

EPA disagrees that additional studies need to be conducted to provide additional information to further constrain certain model parameters, and believes that the results of the studies conducted over the last several years provide sufficient information to meet the goals of the modeling study. In addition, because of the elevated risks documented in the human health and ecological risk assessments, EPA believes it is prudent to move forward with the modeling study rather than delay the project further to conduct additional studies.

An important aspect of model calibration is the reduction of data to achieve consistency with the model inputs, parameters, state variables, etc. This involves spatial and temporal averaging, normalization, and sometimes more involved transformations. There was much data to reduce in this project, and for the most part the modeling team did a masterful job. In several instances, however, the data were not appropriately reduced. Two significant examples, which were problematic for the peer reviewers, were the PCB concentrations in the sediment (used to initialize the sediment bed) and in fish (used to confirm the bioaccumulation simulation). In both cases, the individual data were inappropriately compared to aggregated/averaged quantities in the models. Much confusion resulted regarding what the models were intended to predict and whether the residuals were indicative of model bias.

RESPONSE O-DE-9:

The Reviewer is correct that data reductions were necessary to provide appropriate data linkages among models, to convert input parameters to correct units, and to aggregate data to appropriate spatial/temporal scales. EPA believes that the appropriate data reductions were conducted during model simulations and model linkages, given the objectives of the modeling study (see response to General Issue 1). EPA acknowledges, however, that some confusion was evident in the interpretation of graphical depictions of data analyses, particularly from the presentation of individual sample concentrations plotted against central tendencies from the model simulations.

As noted in the response to General Issue 1, the purpose of the modeling study is to simulate the fate and bioaccumulation of PCBs, with the focus on <u>average concentrations</u> of PCBs in various abiotic and biotic compartments. To provide the most appropriate comparisons of model simulations to field observations, EPA agrees that the individual field observations should be represented by a measure of central tendency and compared to the simulated averages. EPA will address this in the Model Validation Report by: (1) including such comparisons in scatter-plots of simulated versus measured concentrations; (2) clarifying when differences in data reduction methods may have significance for interpretation of figures, and (3) providing a more informative footnote for each figure.

During the Peer Review deliberations, some specific graphs were discussed at length, and issues were raised with respect to: (1) the degree to which data reduction was applied; (2) the comparability of PCB data distributions to central tendency estimates/simulations; and (3) whether the graphs indicated model bias, the variability of individual measurements, or both. Examples are discussed below to clarify some of the specific issues raised:

- “Binning” of Sediment PCBs – Figures 5-5 and 5-6 of the Model Calibration Report summarized PCB concentration data used to initialize the sediment bed. Figure 5-5 plotted measured concentrations in individual sediment samples, whereas Figure 5-6 plotted the central tendencies of the data used to initialize the sediment bed. As the Reviewer notes, the EFDC model simulates only the averaged concentrations in the sediment. In using average concentrations in sediment, EPA has implicitly assumed that the average concentrations in biota will be driven by the average concentrations in abiotic exposure media, which is consistent with the approach used in the risk assessments. Neither explicit modeling of the distribution of sediment PCB concentrations in EFDC nor a full mechanistic understanding of the high variability in sediment PCB concentrations was necessary to satisfy the model objectives (see also responses to General Issues 1 and 5).

- Spatial Scale of Variability in Sediment PCBs – During Panel deliberations, questions were raised with respect to the magnitude of variability observed in sediment PCB concentrations, and whether the graphics shown in the Model

Calibration Report were internally consistent. Specifically, the degree of variability observed in Figure 5-5 was compared against the variability observed in the benthic community sampling program (i.e., 12 replicates collected over a few square meters). The latter indicates the variability over a small patch of relatively homogeneous substrate types, whereas the former indicates variability over larger areas of more variable substrate and hydrodynamic regimes. It is not unexpected that the variability in PCB concentrations associated with heterogeneous substrate across the river profile (i.e., 2 to 3 orders of magnitude) is larger than that observed at specific sampling locations (i.e., 1 to 2 orders of magnitude).

- Fish Tissue Scatterplots – Figures C.3-49 through C.3-52 portray individual PCB concentrations in fish used to evaluate the performance of the bioaccumulation simulation. EPA acknowledges that several of these scatterplots compare individual concentration data (measured) to central tendencies (simulated), resulting in confusion for some Reviewers. The large individual residuals shown on the graphs are indicative of variability in individual fish measurements, but do not indicate substantial bias. When FCM was run in a deterministic mode, average values of all input parameters, including lipid contents, growth rates, dietary patterns, etc., were applied. Individual variations observed in nature for these parameters will result in distributions about the central tendency. The purpose of these graphs was simply to evaluate whether the deterministic simulations systematically under- or overpredicted measured concentrations in the assessment of model bias. The magnitude of variability observed within each combination of species, reach, and fish age (i.e., model precision) is an issue separate from the assessment of model bias.

- Fish Tissue Comparisons of Means – In the Model Calibration Report, several figures compare central tendencies derived from data to central tendencies simulated by the model (e.g., Figure C.3-28, Figure C.3-51, and Figure 6-15). These figures provide the most robust assessment of model bias. In some cases, however, there was only a single sample for a unique combination of age, species, and river reach, and it is not known how well this single data point represents the central tendency. The analysis showed that magnitude of model residuals decreased as the sample size of each group increased (i.e., as the sample means for field data become better indicators of central tendency). This result strongly suggests that the largest residuals are attributable to variability in the data, rather than due to model uncertainty.

In summary, EPA acknowledges a need for improved clarity in the labeling and consistency of "model performance" graphical summaries. Future graphs will be clearly footnoted with an explanation of what is being depicted.

In general, calibration of the models was thorough, if limited by the relatively short time scale of the observations, the previously-mentioned lack of comparison to erosion and deposition measurements, and the omission of a number of processes (e.g., bank erosion) that appear to be significant to the mass balance. The calibrated suspended solids, water column total PCB, and fish PCB concentration appear to be reasonably accurate and unbiased. Many specific comments regarding the calibration are offered below. The major problem with the calibration as reported is that these are short time scale simulations, which are not sensitive to important features of the sediment and contaminant transport models.

RESPONSE O-DE-10:

Please refer to the response to General Issue 2.

Overall, my sense is that the calibration report is an interim deliverable meant to satisfy the timetable of the Consent Decree. Many aspects of the calibration will necessarily be revisited, once a longer-term simulation is constructed and tested.

RESPONSE O-DE-11:

EPA does not view the Model Calibration Report as an interim deliverable produced to satisfy a timetable; it reflects the plan specified in the MFD and Modeling Study Quality Assurance Project Plan (QAPP) (WESTON, 2000). The Model Calibration Report was produced at this time to demonstrate model performance in simulating conditions observed in the data for the time period for which the most intensive data collection effort was conducted and to receive input from the Panel before conducting more resource-intensive efforts. The calibration timeframe was proposed in the original MFD. Please refer also to the response to General Issue 2.

Reference:

WESTON (Roy F. Weston, Inc.). 2000. *Quality Assurance Project Plan: Modeling Study of PCB Contamination in the Housatonic River.* Prepared for U.S. Army Corps of Engineers and U.S. Environmental Protection Agency.

The other efforts documented in the Calibration report (sensitivity analysis and uncertainty analysis) seem redundant at this point, given that much of the model calibration is unfinished. The sensitivity analyses of the models are thorough and informative, however.

RESPONSE O-DE-12:

The sensitivity analyses presented in the Model Calibration Report provided important insights to the modeling team regarding model behavior, focusing further calibration efforts. Therefore, EPA believes it was appropriate to conduct and present the sensitivity analyses as part of the Model Calibration Report.

Only the food chain model has been subjected to rigorous uncertainty analysis. I sense it may be the only one of the models for which this can be practically accomplished.

RESPONSE O-DE-13:

Please refer to the response to General Issue 11.

Wilbert Lick:

Many of my concerns have to do with the modeling framework and data needs. The panel addressed this topic in 2001. However, at that time, the modeling framework was quite general, very ambitious, and had little detail. The framework and model details have changed considerably since then. Because of this, in answering the questions on Modeling Calibration, some preliminary comments on the modeling framework are necessary. Since Question 3 is closest to the concerns about the modeling framework, most of my preliminary comments are included as introductory material to Question 3.

RESPONSE O-WL-1:

EPA acknowledges that the modeling framework has been substantially refined and revised since the first Modeling Framework Design document and the Peer Review of that document. Many of the changes implemented since that time were made in response to the comments of the Peer Review Panel.

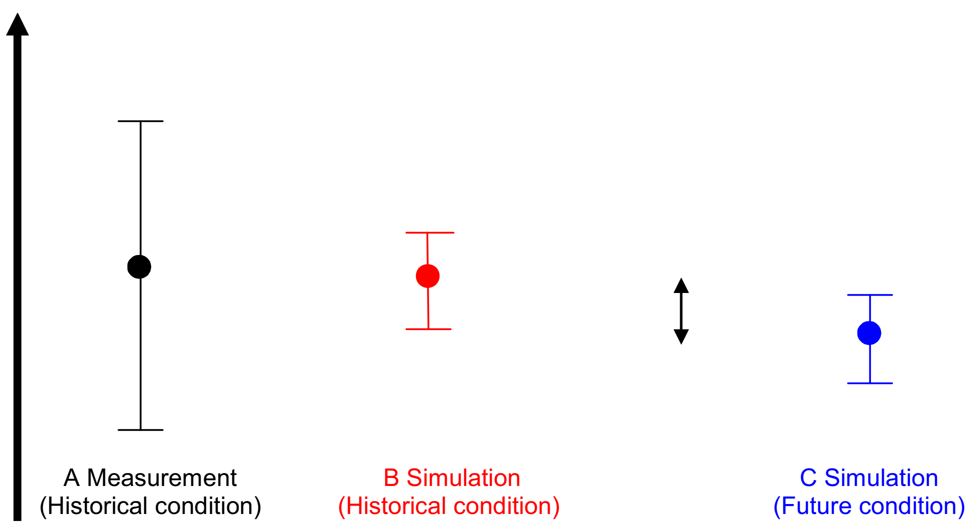
SPECIFIC COMMENTS

Comparison of Model Predictions with Data

1. Are the comparisons of the model predictions with empirical data sufficient to evaluate the capability of the model on the relevant spatial and temporal scales?

E. Adams:

This is really the overarching question associated with model calibration: are the models good enough? But the question has to be asked in the framework of what the models will be asked to do, which is to evaluate remedial alternatives. (Problem identification is usually the first step in model evaluation; see, e.g., Ditmars et al., 1987.) The following sketch and discussion provide a framework for addressing this issue in general terms. Supplemental information is provided in the answers to the remaining five questions.



The black symbols on the left (A) represent, qualitatively, the range in space and time of measured state variables. The models have many state variables, but I will focus on three: flow rate; sediment concentration (bed load and suspended); and PCB concentrations (in the sediment, water column and fish). Some of the variables, such as sediment PCB concentrations, have tremendous variability as implied by the wide brackets. The mean is indicated by the circle. The red symbols in the center (B) represent the calibrated models' prediction of the same variables, while the blue symbols on the right (C) indicate the models' simulation of the same variables under a future remediation scenario.

We are being asked if, following calibration, the models are "good enough". In model calibration, this traditionally means asking how well the simulations (B) match the historical data (A) with respect to the mean, the variance, etc. However, it is more relevant to ask how confident we are that the models will correctly predict the environmental effectiveness of various remediation options, which we are told include removing contaminated sediments by dredging and/or dry excavation, burying them through in situ capping above and/or below water, and natural recovery.

RESPONSE 1-EA-1:

Please refer to the response to General Issue 1.

Evaluation of future options introduces the following questions concerning model skill: how well does the model predict absolute mean output (an absolute measure indicated by the blue circle in C), how well does the model predict relative mean output (i.e., change from existing conditions, indicated by the black arrow representing the difference between the red circle of B and the blue circle of C), and how well does the model simulate the range in output (both in absolute terms and relative to the existing condition). In general a model will perform *better* in a relative assessment, than in a calibration, because model errors tend to cancel. Conversely a model may perform *worse* in an absolute assessment, than in a calibration, because different data and processes (reflecting remediation options) are involved. Of course we don't have measurements of future conditions with which to compare model results so we have to make inferences. This is done qualitatively below for the various types of model output.

Flow rates are primarily an output of HSPF and measured inputs, as filtered by river hydrodynamics (EFDC). HSPF has been around for a long time, its developers have had lots of experience with it, and the available data for the Housatonic seems to be on par with (or better than) what is typically available for other sites. There are a lot of semi-empirical parameters that can be adjusted to achieve a good fit and the fits displayed seem generally acceptable with respect to both mean and variance. (Some errors come when simulating storm events, but this is due to the difficult of getting storms right—both magnitude and timing—with only one hourly rain gauge in the watershed. But this should not be a problem in a statistical sense: if sediment-laden PCBs are being eroded due to a storm, we aren't concerned about the exact timing of the storm.) Because the proposed remediation measures should not significantly affect flows there is no reason to believe the model will not be able to adequately simulate flows under future conditions.

RESPONSE 1-EA-2:

EPA agrees with the Reviewer's assessment that the HSPF watershed model is adequately calibrated to achieve the goals of the modeling study; however, additional calibration results for the Phase 2 period will be presented in the Model Validation Report.

Sediment transport is performed by EFDC (with inputs from HSPF) and results from the processes of erosion, bed load transport (primarily of coarser size fractions), resuspension (primarily of finer size fractions) and deposition. There are no direct field measurements of bed load transport, but the model has been calibrated to produce reasonable agreement with field measurements of mean total suspended solids (TSS) suggesting that the net effect of the contributing processes is satisfactory, at least with respect to the mean.

RESPONSE 1-EA-3:

The Reviewer is not correct that there were no direct measurements of bed load transport. EPA conducted a bed load study that included three sampling events. The study is presented and discussed in Section 4.4.1.2 of the final MFD. Comparison of the results of the bed load study to simulations produced by the sediment transport model are summarized in Volume 1 (Section 4.2.3.1.3) of the Model Calibration Report, and discussed in detail in Appendix B (Section B.3.2.3.2). These comparisons indicate reasonable agreement between the simulated bed load and the data. EPA agrees with the Reviewer's assessment that the agreement between measured and simulated TSS is also reasonable.

The agreement on the variance would best be seen using an equivalence plot of TSS (graph of measured TSS on the vertical axis and simulated TSS on the horizontal axis, with a 45 degree line indicating perfect agreement). Apparently such plots have not been generated, but I suspect that they would show that the model under predicts the variance in TSS for two reasons: First, TSS is a function of erosion, which depends on shear stress to a power n. Shear stress, in turn, depends on the local velocity squared, making erosion dependent on velocity raised to the power of 2n. Because EFDC uses approximately one grid cell per channel width, it can only output channel average velocities. Hence velocity extremes associated with lateral variations in channel depth and meandering are ignored.

RESPONSE 1-EA-4:

Please refer to the responses to General Issues 12 (equivalence plots will be provided in the Model Validation Report) and 4 (regarding the number of grid cells across the channel).

Furthermore, the calibrated values of n are approximately one, whereas W. Lick points out that the literature suggests the value should be more like two.

RESPONSE 1-EA-5:

Please refer to the response to General Issue 7.

Hence a model calibrated to produce the right amount of resuspension on average will likely smooth out the extremes, in particular under predicting resuspension during high-energy events. It would likewise be expected to under predict the extremes in bed load transport. The under prediction of both bed load and suspended load transport would cause the model to underestimate the potential for natural remediation, while the underestimate of bed load transport could cause the model to underestimate the threat of cap erosion if instream capping were to be considered for remediation. I recommend that measured and predicted TSS be plotted on an "equivalence plot" to assess the magnitude of this problem and, if significant, the erosion model be recalibrated. Of course this means that deposition and possibly other processes would also need to be recalibrated.

RESPONSE 1-EA-6:

The equivalence plots recommended by the Reviewer were developed and evaluated. These plots indicated that there was no need to recalibrate the erosion model. Please refer also to the response to General Issue 12.

PCB concentrations are predicted using both EFDC and the Food Chain Model (FCM). EFDC has been calibrated to produce good agreement with measured average PCB concentrations in the water column. This means that the net effect of several exchange processes is in balance (though it doesn't say anything about the individual processes themselves). Over a short calibration period not much change can be expected in the sediment PCB concentrations so it is difficult to assess calibration here.

RESPONSE 1-EA-7:

EPA agrees that while the simulated concentrations are in reasonable agreement with the data, there is little change in sediment concentrations in the PSA during the period of model calibration. This issue may persist through both Phase 2 of Calibration, and in Validation, because of the lack of clear temporal trends in the sediment data over the period of record. Please refer also to the response to General Issue 2.

The FCM seems to be doing a good job of reproducing at least mean concentrations in fish and EPA claims that this is all they care about.

RESPONSE 1-EA-8:

The Reviewer is correct that EPA is most concerned about the average concentration in fish over time. The model performance targets specified in the Modeling Framework Design and the Modeling Study QAPP apply to mean concentrations within each unique combination of river reach, fish species, and fish age. The deterministic application of FCM does not simulate distributions of individual PCB concentration data. EPA believes that the reproduction of mean concentrations by FCM is adequate for application of the model to evaluation of alternate remediation scenarios. Please refer also to the response to General Issue 1.

Remediation may result in lower sediment PCB concentrations (e.g. if some of the PCBs are removed) or a redistribution of PCB mass (e.g., if the PCBs are sequestered under a cap). To the extent that the sediment-water exchange processes, and the biological uptake, are linear with respect to concentration, good agreement between model and data under existing conditions should imply good ability to predict future conditions when the concentrations are lowered due to remediation. To the extent that the remediation removes a fixed fraction of both fine and coarse-grained sediments (and their associated PCBs), which seems reasonable for a dredging scenario, the assumption of linearity is reasonable. (See discussion in the following paragraph.)

RESPONSE 1-EA-9:

Although it is premature to focus on specific remedial alternatives given the process specified in the Consent Decree, the exchange and uptake processes identified by the Reviewer are adequately approximated with the assumption of linearity with respect to PCB concentrations. EPA agrees that demonstration of the ability of the model to predict current conditions provides confidence in the ability of the model to predict future conditions following potential remediation, including natural recovery. Indeed, the concept of model calibration is based on just such an assumption.

I agree with the concern expressed by QEA/GE that the depth of PCB bioavailability within the sediments (6 inches) is too great. Since PCB concentrations are currently fairly well mixed in the upper 6 inches, this does not affect the ability to simulate uptake presently, but it would affect the ability to predict uptake under a future scenario in which the PCBs were buried under cleaner sediments (either by application of a cap, or over time by natural processes).

RESPONSE 1-EA-10:

EPA agrees that the data currently do not indicate strong gradients in PCB concentrations, or other sediment parameters, in the top 6 inches of the sediment bed. This vertical homogeneity reflects biological and physical processes in the PSA that affect the way sediment is deposited and resuspended.

In response to this and other comments received from the Panel, EPA is reevaluating the use of a 6-inch well-mixed layer in the model and is evaluating the implementation of a thinner layer. EPA agrees with the Reviewer that a 6-inch layer may not represent the most appropriate depth for evaluating remedial alternatives throughout the PSA. Please refer also to the response to General Issue 6 for a detailed discussion of the depth of the bioavailable layer.

During the Peer Review Meeting much was said about the tremendous spatial variability in sediment PCB concentrations over space scales of order one meter and the fact that the model can not reproduce this variability. The failure of the model to pick this up should not be considered model error, *per se*, but simply unresolved variability in model input and output (sediment bed concentration distributions).

RESPONSE 1-EA-11:

EPA agrees that it is not possible or necessary for the model to reproduce the observed small-scale variability in PCB concentrations in PSA sediment, and EPA also agrees that the lack of representation of this micro-scale variability is not model error. In addition, EPA believes that the lack of micro-scale modeling (or mechanistic explanation for the observed heterogeneity) is not a significant model limitation, given the objectives of the modeling effort (refer to the response to General Issue 1). The knowledge that this variability exists, however, will contribute to understanding some of the uncertainty in model predictions.

This variability is real and most likely reflects the stochastic method in which the PCBs were introduced in the first place.

RESPONSE 1-EA-12:

EPA agrees with the Reviewer's comment.

We cannot expect the model to predict this variability and the fact that the model averages concentration over relatively large grid cells is not a problem (with the mean) unless sediment-water exchange of PCBs varies non-linearly with concentration.

RESPONSE 1-EA-13:

EPA agrees with the Reviewer that the model cannot be expected to predict the variability observed in the data and that the use of mean concentrations does not pose a problem because sediment-water exchange of PCBs is represented in the model by dissolved and particulate exchange that varies linearly with concentration.

The exchange of dissolved PCBs is computed based on a surface mass-transfer coefficient and the concentration gradient between the dissolved PCB concentration in the surface water and that in the pore water. Because the pore water concentrations are much greater than the dissolved concentrations in the surface water, the dissolved transport varies linearly with sediment PCB concentration. The model, therefore, simulates the correct average flux of PCBs between the water column and sediment compartments on the basis of the average concentrations.

The exchange of particulate PCBs is calculated based on PCBs associated with each of the four solids classes (i.e., one cohesive and three non-cohesive classes). In each grid cell, the PCB concentrations on the different solids classes are calculated based on the total PCB (tPCB) concentration, the fraction of the total sediment in each solids class, and the different organic fractions on each solids class. Based on these calculations, the exchange of particulate PCBs is also linear with concentration.

Of course, we can not expect the model to tell us anything about the future variance of sediment bed concentrations, and to the extent this is important we should rely on the observed variability. The PCBs have been in the sediments for several decades, and to a first approximation the variability expected in the next decade or two (presumably our focus) will not be very much different from the variability observed historically.

RESPONSE 1-EA-14:

EPA agrees with the Reviewer's comment.

Back to the mean, sediment-water exchange *would* be expected to vary non-linearly with sediment concentration to the extent that the PCBs are associated preferentially with finer sediments that are more easily eroded. The model could be getting the flux of (primarily fine grained) sediments correct, but would be assigning an average PCB concentration to these sediments and hence underestimating the flux of PCBs. As a result I suspect the effect of averaging (sub-grid scale variability in sediment PCB concentrations) results in an underestimate of PCB flux during periods of high resuspension, for both existing and future conditions.

RESPONSE 1-EA-15:

Preferential sorption of PCBs to finer sediment is recognized and accounted for in the modeling approach. Formulations are used for erosion that account for bed composition, which eliminates the potential problem discussed by the Reviewer.

The distribution of PCBs among different solids size classes within individual sediment samples was investigated as part of the EPA Rest of River Study. Sediment solids were separated by grain size, and PCB and TOC were measured on each size fraction. Carbon normalization reduces the variability in PCB concentrations measured on the different size classes (see Model Calibration Report Appendix B, Figures B.4-22 through B.4-25).

In the model, the PCB concentrations on the different solids classes are calculated based on the tPCB concentration, the fraction of the total sediment in each solids class, and the different organic fractions on each solids class. The size fraction composition of the solids eroded from the bed is calculated based on the composition of the bed material and recognition that the bottom shear stress required to erode non-cohesive particles increases with increasing particle size. Resuspension of cohesive solids is calculated based on site-specific erosion data.

To assess this effect, the correlation of PCB concentration and sediment type should be checked in available measurements and, to the extent possible in prediction. (I realize that model output is averaged over spatial scales that include a range of sediment type.)

RESPONSE 1-EA-16:

Figures 4-16 and 4-17 of the RFI (BBL and QEA, 2003) illustrate the relationship observed in the data. These figures indicate that there is only a weak relationship (i.e., generally low r^2 values) in the PSA between sediment properties and PCB concentrations. Similar figures will be included in the Model Validation Report to address this comment with respect to the model results.

Also, although it would be a major change at this point, I wonder if it wouldn't be better to have the model formulated to predict sediment PCB concentrations simply as a function of sediment type (e.g., coarse, medium and fine) with only very coarse longitudinal discretization (say by reach 5A, 5B, etc.) This would result in much less model output (by two orders of magnitude), making the calculations more tractable, and the output would be more environmentally relevant: we don't care which 20 m cell within a reach a fish is in when it feeds in contaminated sediments, but merely the likelihood that the particular sediments will actually be contaminated—and to what extent.

RESPONSE 1-EA-17:

EPA agrees in principle with the Reviewer that coarse longitudinal binning of simulated contaminant concentrations is sufficient to achieve the goals of the modeling study; in fact, FCM operates on exactly such a scale, and EFDC output is averaged over subreach bins before it is passed to FCM. Simply relating contaminant concentrations to sediment type is not possible, however, because sediment characteristics are only one of many factors involved in contaminant transport and fate, as indicated by the sometimes weak relationships between sediment type and PCB concentrations in the PSA. The additional factors controlling contaminant transport and fate in the PSA are discussed in detail in Section 4 of the MFD, and in part determined the model selection and other aspects of the modeling framework. Adequate simulation of all the factors determined to be of importance in the PSA is necessary for satisfactory prediction of PCB fate, even if the output is subsequently integrated to larger spatial and temporal scales.

W. Frank Bohlen:

The calibration period extends from May 1, 1999 to June 30, 2000. Although this was a period that allowed for the sampling of a range of average ambient and aperiodic storm conditions it is, from the standpoint of sediment/contaminant transport in the Housatonic River study area, a very short period of time. Erosion processes affecting the side banks and the associated channel migration, sediment deposition in the backwaters and Woods Pond and many of the transport processes affecting the floodplains operate on time scales long compared to the calibration period. As a result the comparisons conducted over the calibration period provide only limited indication of the model's ability to accurately predict longterm change. There is some indication that this is recognized by the model developers and will be addressed during the verification phase. Such use of the verification phase for calibration purposes is not recommended.

RESPONSE 1-FB-1:

The time period used for model calibration was proposed in the MFD; only a single comment from the Reviewers was received on the appropriateness of the approach. In response to Reviewer comments on the Model Calibration Report, EPA is revising the approach to calibration and validation; please refer to the response to General Issue 2.

If 14 months is too short what might be an adequate calibration period? This is a question that would benefit from some amount of discussion by those most familiar with the study area. At present the reports provide relatively little discussion of the reasoning that lead to the selection of the 14 month period. Although I've been trying to encourage a shorter report this is a subject that would benefit from additional discussion. My brief review of available data detailing sedimentation in the study area as well as contaminant concentrations suggests that a five year period of calibration would result in a more robust test of model capability and complement the longer term validation runs.

RESPONSE 1-FB-2:

Please refer to the response to General Issue 2.

Douglas Endicott:

I can't answer this question with a "yes" or "no" for the calibration of the EFDC model, because I do not believe it is being applied on spatial and temporal scales that are relevant to the PCB contamination problem in the PSA and its remedy. I will address the issues of relevant spatial and temporal scales in the following paragraphs.

RESPONSE 1-DE-1:

Please refer to the responses to General Issues 2 and 4.

In the case of the food chain model, it appears that comparisons made between predictions and observations are adequate, except in reach 5D where no fish were sampled.

RESPONSE 1-DE-2:

EPA agrees that the model performance measures for FCM were adequate. The implications of the lack of fish sampling in Reach 5D backwaters are discussed above in Response O-DE-7.

I have a fundamental objection with the application of EFDC as a predominantly 1-dimensional model to simulate hydrodynamics, sediment transport and PCB transport in the PSA. Accurate simulations of velocities, shear stresses, erosion and deposition patterns, and streambank undercutting/erosion are only possible if significant lateral variations are resolved in the model. At least 3 lateral segments should be used in the main river channel, and should consider river features such as bathymetric profiles. This lateral segmentation should also be used in the sediment bed, with initial sediment conditions recalculated from data on the basis of this segmentation. Fortunately, EFDC is a 3-dimensional model so it should be able to accommodate this additional resolution.

RESPONSE 1-DE-3:

Please refer to the response to General Issue 4.

The temporal scale of EFDC calibration is not particularly relevant to the PCB contamination problem in the PSA and its remedy. It should be noted that "calibration" in the context of this project has come to mean calibration of short-term (daily to seasonal) changes in model state variables. "Validation" now includes the calibration of long-term (annual to decadal) changes, which are the interesting changes in terms of managing toxic chemicals and making decisions about remedial alternatives. In other words, at this juncture we are unable to evaluate the model's capabilities in terms of its intended application. In order to fully evaluate the model's capabilities, it must be applied to a significantly longer simulation period (i.e., 10 years) and compared to data over this longer duration. This limitation also has an impact on how thoroughly we can address questions 4, 5 and 6 below.

RESPONSE 1-DE-4:

Please refer to the response to General Issue 2.

Regarding the calibration of sediment transport in EFDC, there is too much emphasis on matching suspended solids concentrations, and not enough on scour and deposition in terms of changing sediment bed elevation. The calibration of suspended solids only fixes the magnitude of net settling/resuspension in the model, while the individual fluxes may or may not be correct. Confirming the change in sediment bed elevation can be much more revealing.

RESPONSE 1-DE-5:

EPA does not agree that too much emphasis was placed on suspended solids in the model calibration. Transport of suspended solids is the primary transport mechanism for PCBs in the system. However, it is true that the change in sediment bed elevation is also an important fate-controlling factor for PCBs in the system because, in depositional reaches at least, the rate of change in sediment bed elevation (i.e., the net sedimentation rate), in combination with the well-mixed layer depth and the concentration of PCBs associated with the depositing solids, control the rate of long-term recovery of the system. The model currently computes a net sedimentation rate that ranges from 0.07 to 0.24 cm/year on a reach-average basis. While this may be low in comparison to the limited information on net sedimentation rates for the system, it is necessary that the comparison be made using results from a longer term simulation period, as will be available at the conclusion of the model validation phase of the investigations.

The most robust data documenting changes in sediment bed elevation are from the river channel surveys collected at nine transect locations from September 2000 to September 2003 (final MFD, Appendix H.2). An additional survey was performed at these locations in July 2005, and the locations were resurveyed following the high-flow event of October 2005. From examination of these results, and from additional anecdotal information and observations of channel bed elevation changes following storm events, it is apparent that scour and deposition in the PSA are phenomena that are extremely variable spatially, are occurring over small spatial and temporal scales, and appear to have a high degree of stochasticity. In combination, these factors make scour and deposition a difficult parameter for the model to reproduce exactly because the net scour or deposition in an individual model channel cell at any time step must reflect the net result of a number of individual scour and depositional areas, and because some parameters affecting scour and deposition (such as the presence of woody debris) cannot be represented in the model. Because the model is not intended or expected to represent small-scale scour and deposition, it is difficult to use the site-specific data to evaluate model calibration, except as an indicator of typical behavior over areas of the river.

In that regard, I have some trouble reconciling the 3 to 4 cm maximum change in sediment bed elevation predicted by the sediment transport model for the 14-month calibration period, with the several-foot change based on bathymetric transects and stratigraphic analysis. I am afraid the calibration results fail to demonstrate that the approach taken to calibrate settling and resuspension fluxes works.

RESPONSE 1-DE-6:

The difference between the model results and the observations referred to by the Reviewer is a result of several factors. First, no large-scale events of sufficient magnitude to redistribute several feet of sediment occurred during the Phase 1 model calibration period. Additionally, the model cells are typically about 20 meters long, while the bathymetric transects indicating marked erosion of the cross-section represent point observations along the longitudinal axis of the river. Thus, the model segments provide an estimate of reach averages and would be expected to smooth out finer scale changes in bathymetry that may occur. For this same reason it is important that the model correctly evaluate the water column TSS level over time because the TSS level reflects the cumulative impact over distance of the variable fine-scale changes in channel bathymetry that have occurred upstream of the point of sample collection.

The chronological order of the lines in the plots of changes in bathymetry at the cross-sections is another factor that needs to be considered because the beginning and end of storm profiles do not necessarily correspond to the maximum range in elevation that occurs over the course of a storm. Thus, the net change in elevation during the storm may be less than is suggested by the range in elevation that is reported throughout the storm. The comparison of these transects with the model must necessarily be performed during the validation phase of the analysis because the cross sections were not measured within the 14-month calibration period. Please refer also to the response to General Issue 12.

How do we know that the highly-nonlinear parameters describing things like cohesive sediment erosion, flocculation, and deposition can be determined by averaging data collected at a number of sites having different sediment properties? To a certain extent, the answer is obtained by running the model for a relatively long duration, and examining the results for anomalies in terms of the magnitude and pattern of sediment bed change.

RESPONSE 1-DE-7:

Please refer to the response to General Issue 12.

In some other river systems, errors in deposition and erosion fluxes are revealed during calibration of water column PCB concentrations, because the concentration gradient between sediment and water amplifies the error. In the PSA this will not work so well, because there is only a small gradient between suspended and bedded particulate PCB concentrations. The lack of gradient makes it relatively more difficult to tell if the model is grossly in error.

RESPONSE 1-DE-8:

It is acknowledged that the water column and bed sediment particulate PCB concentrations in the Housatonic River may not be as different as in some other systems, when compared on a whole sediment basis. However, the difference is actually amplified considerably when the PCBs associated with the fine sediment fraction are compared to the water column particulate PCBs. This is because the fine sediment fraction is higher in organic carbon content and thus, has a relatively high dry weight PCB concentration as well (see Figures B.4-22 through B.4-25 in the Model Calibration Report).

Because the fine sediment fraction is the primary fraction that contributes to the resuspension flux, the PCB concentration associated with the fines serves as a useful tracer that places a constraint on the flux of solids to the water column. The solids flux between the water and sediment needs to be such that water column TSS levels are correctly simulated, although it is acknowledged that correct simulation of TSS can be achieved by a combination of incorrect settling and resuspension rates and therefore, is not, in itself, sufficient for demonstrating calibration. The spatial profile of water column PCB concentrations does, however, provide an additional constraint on the magnitude of the settling and resuspension fluxes of PCBs to the water column.

In combination, the reasonableness of these comparisons (for both solids and PCBs) indicates that the calibrated model is not "grossly in error." The response to General Issue 7, "Resuspension and Deposition," presents additional discussion of the constraint imposed by the gradient between particulate PCB concentrations in the water column, and on cohesive solids in the sediment bed.

The parameterized value of the mass transfer coefficient for pore water diffusion (K_f) is very high in comparison to most values I can find in the literature. K_f is being calibrated to reproduce the observed increase in water column PCB concentrations under low-flow conditions. In other words, all of the increase in water column PCB concentrations is being allocated to this mechanism. Whether the K_f calibration is correct depends upon this assumption. It would be most desirable to somehow independently confirm this value, either via measurement or by ruling out other potential PCB sources.

RESPONSE 1-DE-9:

The value of K_f (1.5 cm/d), calibrated to Housatonic River data, is at the low end of the range of values summarized by Thibodeaux and others (Thibodeaux and Bierman, 2003; Thibodeaux et al., 2002). These authors noted intra-annual variations in K_f values in the Hudson, Grasse, and Fox Rivers, with lower values (3 to 10 cm/d) in the winter and higher values (20 to 40 cm/d) in the summer. It is noted that the minimum value of 0.2 cm/d for the Fox River above DePere Dam comes from a model (Velleux and Endicott, 1994) that includes "background" resuspension at low flow conditions, which the authors acknowledge "substantially influences water column PCB concentrations, but has little impact on solids concentrations." In addition, please refer to the response to General Issue 8.

References:

Thibodeaux, L.J. and V.J. Bierman. 2003. The bioturbation-driven chemical release process. *Environmental Science and Technology* 37:252A-258A.

Thibodeaux, L.J., G.A. Canin, and M. Cain. 2002. Recent advances in our understanding of sediment-to-water contamination fluxes: the soluble release fraction. Keynote presentation at the 5th International Symposium on Sediment Quality Assessment of Aquatic Ecosystems and Public Health, Chicago, IL. October 16-18, 2002.

Velleux, M. and D. Endicott. 1994. Development of a mass balance model for estimating PCB export from the Lower Fox River to Green Bay. *J. Great Lakes Res.* 20(2):416-434.

Rare flood events

The calibration report shows that EFDC is capable of predicting the extent of flooding. However, this is not an *impact* per se. The impact of concern is the remobilization of significant quantities of previously in-place pollutants.

RESPONSE 1-DE-10:

EPA understands that flooding, in and of itself, does not represent an "impact" and that the goal of the modeling study is to be able to predict contaminant transport and fate under different remedial options and natural recovery. The comparison of historical aerial photographs taken during flooding events to model predictions of the extent of flooding during the same events was conducted simply as another means of comparing the reasonableness of model output during an extreme event. The model was able to reproduce both the general extent and certain important features of the flooding observed during the events tested, thereby providing additional confidence in the calibration of the hydrodynamic submodel.

Sediment transport models like SEDZL and now EFDC are being used in a growing number of river systems to predict bed erosion under event conditions. However, the state of the art still requires extensive site-specific and model process-specific data. As far as I know, confirmation of model predictions under extreme events must still be demonstrated on a site-specific basis. Based on the calibration report, this confirmation is lacking in the Housatonic River.

RESPONSE 1-DE-11:

EPA made every attempt to collect data during an extreme event until just recently; however, because of forecasting and/or logistical issues such efforts were unsuccessful. EPA does not believe it is worthwhile to delay the modeling study to wait to gather data from such an event. In the absence of data, any available information that existed or could easily be collected (such as extent of flooding) for a large event was used to the extent possible to test model predictions.

EPA's field program included several elements specifically designed to provide a basis for parameterizing model processes (e.g., Sedflume erosion measurements), and evaluating model predictions (e.g., bed load sampling and storm event sampling). While it would have been fortunate to have obtained data at higher-flow conditions, and plans were in place to monitor such an event, it did not happen within the necessary timeframe. EPA has acknowledged that the longer-term modeling, which is underway, will provide a more rigorous test of the model parameterization. The longer-term modeling will cover time periods when additional data collection efforts were conducted, such as the resurveyed cross-sections, providing additional lines of evidence that can be included in the evaluation.

It is unclear whether the EFDC scour and deposition predictions are reasonable at very high flow rates, and the comparisons to data are problematic due to the spatial resolution of the model sediment bed. Although bathymetric data showing scour and deposition patterns were collected and reported at a number of transects in the PSA, this data has apparently not been used to confirm model predictions as it has in other river systems (e.g., Gailani et al., 1996.).

RESPONSE 1-DE-12:

Cross-sectional surveys measuring scour and deposition at a number of transects within the PSA were obtained during a time period outside the May 1999 to June 2000 calibration period. These data will be used in qualitative comparisons with the longer-term modeling results, although as discussed in Response 1-DE-6, erosion and deposition patterns described by the resurveyed cross-sections reflect small-scale effects that may not be represented by average conditions over the length of a 20-m cell. Please refer to the response to General Issue 5 for additional discussion of the spatial resolution of the sediment bed in the model.

Discriminating between water-related and sediment-related sources of PCBs to fish and other biota

The transport/fate and food chain models address PCB bioaccumulation via both pelagic and benthic exposure routes. In principle, it would be simple to apply the models to discriminating between water-related and sediment-related sources of PCBs to fish and other biota. I have done this in other applications by running the model with two chemical state variables, one for the chemical initialized in the sediment bed and a second for the chemical originating from water column sources. Since the models are both linear with respect to chemical concentrations, the simulation can be decomposed in this manner to explicitly show the proportion of PCB body burden in different species and reaches contributed by PCB exposure originating in the sediment bed versus the water column. If this discrimination is important, the models should be rerun to make this diagnosis.

RESPONSE 1-DE-13:

EPA agrees that the bioaccumulation model kinetics can be partitioned to discriminate between PCB uptake attributable to benthic versus pelagic sources. The model can also discriminate among other combinations of PCB sources (e.g., uptake derived from dietary uptake versus that derived from gill exchange). During model development and calibration, such tests were applied as a quality assurance (QA) procedure to ensure that model output was reasonable. EPA will consider documenting the results of similar tests applied during Phase 2 Calibration and/or Validation.

Marcelo H. Garcia:

Comparisons between model predictions and observations have been made with some apparent success for a relatively short period of time (i.e. several months). In addition, the smallest spatial scale resolved by the model is on the order of the river channel width (i.e. tens of meters). The model seems capable of reproducing observations made during storm events. However, the capability of the model to predict long-term effects (i.e. years) and small-scale processes (i.e. mass transfer at sediment-water interface, bank erosion) remains to be shown.

RESPONSE 1-MG-1:

EPA believes that the addition of a second phase of calibration using a longer period of record and model validation will provide the capability to demonstrate model performance, to the extent that the data record allows. EPA agrees that the duration of the Phase 1 Calibration may not have sufficiently demonstrated the capability of the model to predict small-scale processes, and bank erosion was not included in the Phase 1 Calibration (but will be included in Phase 2). Please refer to the response to General Issue 2.

For example, the model should be capable of predicting sediment depositional patterns as well as PCB distribution in Woods Pond over a time period of decades. The data to test the model is already available in the sedimentary record.

RESPONSE 1-MG-2:

Simulated sedimentation rates for the calibration period were compared to sedimentation rates estimated from analysis of radioisotope-dated cores (see Model Calibration Report, page B.3-51). EPA acknowledges that comparisons of these data with longer-term model simulations will be more meaningful. Those comparisons, and comparisons of PCB distributions in Woods Pond, will be included in the Model Validation Report.

Frank Gobas:

Tables 2.5 and 2.6 demonstrate that *mean* water flow rates are very well predicted over time and space by the **HSPF** model. Despite the good agreement of model predicted and observed *mean* flow rates, daily and monthly scatter plots at Coltsville illustrate that there are considerable variations around the mean flow rates. The variability around the mean is almost one order of magnitude. This variability around the mean value is not represented in the measures used to characterize the quality of model calibration. I recommend that they are added.

RESPONSE 1-FG-1:

Tables 2-5, 2-6, and 2-7 include statistics for annual, monthly, and daily flow comparisons for the calibration period. Moreover, the flow duration curves for both daily and hourly flows (Figures 2-13 through 2-16) demonstrate a very good to excellent representation of the data over the full range of flows observed during the calibration period. This is an overall confirmation that the watershed model represents the hydrologic regime of the Housatonic River watershed, and is not focused on just the mean flows.

The variability in the scatterplots, which is about an order of magnitude for daily flows but is much less for monthly flows, is expected due to the large spatial variability in precipitation in the watershed. Only a single hourly record was available to mimic the rainfall timing across the watershed for the entire calibration period, and was used to distribute daily totals from three other long-term precipitation records (discussed in Section 2.2.2 and Section A.2.2.1). The "scatter" in the scatterplots is not a variation about a mean flow, but more a variation from a "perfect fit" of the model to the data. Perfect fits cannot be expected for watershed models when precipitation, which is the major boundary condition or forcing function, must be estimated from four point observations and extended to the entire 282 square miles of the watershed area. In spite of this necessary extension of the available precipitation data, the weight of evidence summary demonstrates a good to very good overall calibration, and one that meets the QAPP targets.

They it should be considered when applying the model under scenarios where temporal variations as well as maximum and minimum flow rates are important. The *mean* flow estimates of the HSPF model are likely sufficient for addressing the most important management questions such as the response time of contaminant concentrations following remediation options.

RESPONSE 1-FG-2:

Although it is important to adequately simulate mean flows, reproducing the general pattern of magnitude and timing of flow is of primary importance for input to EFDC. EPA believes that the general pattern of flow is also adequately simulated and will achieve the goals of the modeling study.

Table 2-12 and 2-13 illustrate that predicted and observed TSS loading rates are also in good agreement both on a spatial and temporal scale. In comparison to the water flow rates, the TSS model predictions show larger discrepancies between observed and predicted values. Differences of up to 139% are reported. However, the average difference is approximately 10%. The comparison of model predictions and empirical data appears to be sufficient to make estimates of *mean* TSS loads under normal conditions. But again, I recommend that additional detail is provided in the report to better represent the capability of the model to make predictions on spatial and temporal scales.

RESPONSE 1-FG-3:

The model predictions for total suspended solids (TSS) loads show greater differences when compared to data than predictions for flow for a number of reasons:

- The inherent nature of sediment erosion and transport is much more dynamic than stream flow, and is often represented as a power function of flow. Therefore, small to moderate differences in flow rates (i.e., simulated versus measured) are magnified many times when used as the basis for sediment erosion and transport calculations.

- The entire annual TSS load is often transported in a few storm events. If flow differences occur in the simulation of those particular events, the discrepancies for the corresponding TSS loads are often magnified.

- Measurements of TSS are more difficult to collect than flow measurements, and as a result, the data for the Housatonic River are more limited. Consequently, the data used for comparison of annual loads in Tables 2-12 and 2-13 are not actual measurements, but estimates derived from the EPA flux analyses and the RFI report using various flow-sediment relationships. This also leads to greater percent differences than would be expected for the flow simulation.

Note that the greatest percent differences in Table 2-13 are during years with relatively low TSS loadings, and the smallest discrepancies are during the years with high TSS loadings. This is one reason why the mean annual differences are relatively minor compared to the year-to-year differences because the high load years will dominate the calculation of the mean value.

In response to the recommendation from the Reviewer, additional results will be provided in the Model Validation Report to demonstrate the model's capability for simulating TSS loadings.

Tables 2-15 and 2-16 show that differences between observed and predicted water temperatures are very small. These differences are essentially insignificant and the model's capability to predict temperature is very good.

RESPONSE 1-FG-4:

EPA agrees with the Reviewer's assessment.

With regards to the **hydrodynamic** model, empirical observations are to a large degree internalized in the model. The comparison of observed and predicted is therefore not an independent test of the capability of the model.

RESPONSE 1-FG-5:

The forcing functions for the hydrodynamic model are the boundary inputs, which are dominated by the East and West Branches. These inputs were based on data when available, and when data were not available, from estimates developed primarily from relationships between flow at the USGS Coltsville gage and flow upstream of the Confluence. Comparisons between data and model simulations at downstream locations reflect the response of the model's description of the physical system (bathymetry, topography, and friction) and the forcing functions.

EPA believes that these comparisons provide a successful demonstration of the capability of the model. In addition, the ability of the model to reproduce the spatial extent of flooding documented by aerial photographs following Hurricane Bertha support this conclusion. The ability of the model to reproduce the extent of flooding in the floodplain is an important element of the transport of PCBs onto the floodplain.

The model was tested for two extreme events and showed good results. This is promising, but it is not sufficient to conclude that the hydrodynamic model has the capability to predict the hydrodynamics at the relevant spatial and temporal scales.

RESPONSE 1-FG-6:

EPA believes that the spatial scales incorporated in the Phase 1 Calibration were sufficient to achieve the goals of the modeling study. The temporal scale will be addressed in the Phase 2 Calibration and Validation. Please refer to the responses to General Issues 2 and 4.

In my view, it is premature to comment on the capability of the model at this point. The real capability of the model will be revealed in the model validation phase, which will provide a relatively independent test of the capability of the model.

RESPONSE 1-FG-7:

Please refer to the response to General Issue 2.

There is a reasonable data base available to test the **sediment transport** model at two locations (i.e. New Lennox Road and Woods Pond Outlet). Data from other locations (i.e. Holmes Road and Woods Pond Headwater) exist but the sample size is not large. Figures 4-34 and 4.35 illustrate that the sediment transport component of the EFDC model has reasonable central tendency characteristics.

RESPONSE 1-FG-8:

EPA agrees with the Reviewer that Figures 4-34 and 4-35 in the Model Calibration Report indicate the model produces reasonable central tendencies. These figures also show differences between the amount of variability in the model results and variability in data, some of which is attributed to short-time-scale variations in data that the model is not expected or intended to reproduce. Time series plots of TSS data during storm events (e.g., Appendix B, Figures B.3-46 and B.3-47) show small time-scale variations in TSS concentrations of a factor of two or three, scattered, both higher and lower, around the model simulation.

In the case of TSS concentrations measured at New Lenox Road during the September 15-19, 1999 event, fluctuations of more than 100 mg/L in samples collected within the same hour were evident. These fluctuations are not reproduced by the model in part because of the approach used to specify the upstream boundary conditions. Because there is no reason to believe that linear interpolation between rapidly fluctuating concentrations at the boundaries would provide a reliable estimate of concentrations at times between measurements, the boundary conditions were developed as 3-hour moving averages of the storm event monitoring data. Given the unexplained variability at the upstream boundary, this approach is reasonable, even though it contributes to a reduction in the variability in simulated concentrations.

However, there are also significant discrepancies between observed and predicted data. Differences between measured and simulated TSS data show that predicted TSS produce a narrower range of concentrations of TSS concentrations than observed. Also, there appears to be a considerable variability in the measured TSS data at New Lennox Road and Woods Pond Outlet that is not explained by the model.

RESPONSE 1-FG-9:

Please refer to the response to General Issue 10.

In terms of assessing the spatial capabilities of the model, it would be beneficial to have access to more data for model-data comparison but the currently available data sets can be considered adequate as long as the magnitude of the uncertainties are recognized by the model and considered when remedial options.

RESPONSE 1-FG-10:

EPA agrees with the Reviewer that model uncertainty must be recognized and considered in the evaluation of remedial alternatives. The Model Calibration Report included a description of the plan for assessing uncertainty, which was presented to the Peer Review Panel at the Document Overview Meeting. This topic is also discussed in the response to General Issue 11.

I think that the reported analyses can be improved upon by explicitly recognizing the variability among the individual data/prediction comparisons.

RESPONSE 1-FG-11:

Please refer to the response to General Issue 10.

The capabilities of the sediment transport model on a temporal scale are tested over a 14 month period starting in May 1999. A reasonable number of data is available for model calibration. However, Figure 4-25 shows that significant discrepancies exist between observed and predicted TSS concentrations. Therefore, some doubts remain with regards to the temporal capability of the model.

RESPONSE 1-FG-12:

As the Reviewer pointed out in a previous comment (see Comment 1-FG-8), substantially more data are available for New Lenox Road and Woods Pond Outlet (Model Calibration Report Figures B.3-43 and B.3-45, respectively) than for Holmes Road (Figures 4-25 and B.3-42) or Woods Pond Headwaters (Figure B.3-44). The model-data comparisons for New Lenox Road and Woods Pond Outlet show better agreement than for the stations with fewer data. The substantial difference in the number of data points results from the fact that New Lenox Road and Woods Pond Outlet were sampled as part of the storm event monitoring program, but data at Holmes Road and Woods Pond Headwaters were collected primarily as part of GE's monthly monitoring program. The GE samples were collected starting over 20 miles downstream of the PSA and moving upstream, in some cases over more than a single day. Because of this sampling protocol, concentrations at downstream locations can be affected by boundary concentrations that entered the PSA several days before the time of collection of the data used to assign the upstream boundary conditions.

At this point it is unclear whether this quality of agreement between observations and predictions is due to systematic errors in the modeling approach or reflects statistical variability or uncertainty in TSS concentrations. In my view, much value will be added to the modeling effort if in addition to the central tendencies of the model, variability and uncertainty are recognized and explicitly stated.

RESPONSE 1-FG-13:

Please refer to the response to General Issue 10.

Comparisons of predictions of the sediment transport model and empirical data for several storm events are also presented. The agreement between measured and simulated TSS and flow data are with some exceptions are quite reasonable.

Regarding **EFDC**, Figure 5.11 shows the comparison between predicted and observed concentrations of PCBs in pore water. The comparison is quite good, suggesting that the assumption of equilibrium between PCB concentrations in sediments and pore water is justified.

RESPONSE 1-FG-14:

EPA agrees with the Reviewer that the referenced model-data comparisons indicate good agreement.

Figures 5.17 to 5.19 show that the comparison between predicted and observed PCB concentrations in the water column over 14 months at 3 locations. Figures 5-20, and 5.21 illustrate the comparison of observed and predicted PCB concentrations in the water column at various locations in the River. Figure 5.23 illustrates the comparison of PCB concentrations in water column after a storm event. The agreement of the model with the data appears quite reasonable. This is in some contrast to the results depicted in Figure 530, which illustrates a reasonable central tendency of the model in predicting PCB concentrations in the water column, but also considerable discrepancies between model predictions and observations.

RESPONSE 1-FG-15:

The various figures presented in the Model Calibration Report provide alternative ways to compare model results and data. Each has advantages and disadvantages, but collectively they provide what was intended to be an unbiased illustration of the strengths and weaknesses of the calibrated model. Generally, the model does reasonably well with regard to predictions of the central tendency of the data, but individual data points can be either over- or underpredicted at various times, particularly during high-flow events.

The spatial profiles display the cumulative effects on water column PCB concentrations from the interaction between water and sediment as water flows through the PSA. The model does reasonably well at capturing this response. Similarly, the 14-month time-series plots show how the model is able to consistently predict rapid increases in concentrations during runoff events, although the magnitude of the peaks and the detailed timing of the changes in concentration over very short time scales are difficult to discern on these graphs. These comparisons are more clearly displayed on the time-series plots for individual events (e.g., Figures 5-22 and 5-23 and related figures in Appendix B, such as Figure B.4-44). A phase shift of a few hours between the predicted concentrations and the data (e.g., tPCBs at New Lenox Road, Figure B.4-44) is judged to be inconsequential with respect to the evaluation of exposure levels for use in FCM and for the evaluation of remedial alternatives, as long as the longer-term average concentrations are reasonably well represented.

These differences between model results and data are highlighted when displayed on the cross plots of simulated versus measured concentration, such as those shown in Figure 5-30. The same deviations between model and data are displayed in both the time-series plots and the cross plots, but in the cross plot a large difference that results from a slight time shift of as little as 1 or 2 hours appears the same as a clear over- or underprediction of the magnitude of the impact of a runoff event. As a result, the model results appear less favorable relative to comparisons that are made in the context of a storm (i.e., comparisons on a time series plot). The discrepancy resulting from the time shift, however, is of little importance for the intended use of the model.

Additional attention could be devoted to the comparison of observed and predicted concentrations of PCBs in bottom sediments of the river. These data are likely to be very useful in assessing the fate of PCBs in the River.

RESPONSE 1-FG-16:

EPA agrees that the response of PCB concentrations in sediment over time is an important aspect of the model. Changes in sediment concentrations were, as expected, minor over the course of the 14-month Phase 1 Calibration. More attention will be given to comparisons of model and data as the longer-term Phase 2 Calibration work proceeds. Please refer also to the response to General Issue 2.

In terms of the adequacy of the model-data comparisons to evaluate the capability of the EFDC model on the relevant spatial and temporal scales, there appears to be a reasonable amount of data available to evaluate the capability of the model to assess water column transport. The capability of the model to assess some other key aspects of the fate of PCBs, such as long term response times of the PCB concentrations in the River, is not convincingly demonstrated in my view.

RESPONSE 1-FG-17:

The Reviewer is correct that changes in PCB concentrations in sediment over the 14-month calibration period are minimal, and therefore, cannot be demonstrated as part of the calibration. The ability of the model to predict changes in concentration of PCBs in sediment will be further tested over the course of the longer-term Phase 2 Calibration period; however, the data do not show clear trends in PCB concentrations during that time frame as well. Please refer also to the response to General Issue 2.

As for the **bioaccumulation** model, comparisons of model predictions and empirical data on a spatial scale are shown in Figures 6.5 to 6.16. There are additional comparisons presented in Figures 2-34 of attachment C15. The concentrations of tPCBs in sediments and suspended solids show small differences among the reaches 5A to 6. Hence, the model calculations of the tPCBs in biota do not show a strong spatial dependence. As a result, the capability of the model to make spatially explicit estimates could not be fully explored in this study. However, this is not of great importance for the development of the bioaccumulation model as the spatial (and also temporal) differences in concentrations are predominantly determined by other components of the model.

RESPONSE 1-FG-18:

EPA agrees that the model simulations of PCB concentrations in biota do not show a strong spatial dependence due to the relatively flat gradient in PCB concentrations observed in exposure media across the PSA. EPA also agrees that the limitations in the ability to make spatially explicit estimates are not of "great importance." However, EPA has acknowledged the value of conducting activities during model validation that consider PCB exposure levels different from those observed in the PSA.

In conjunction with the model validation, EPA intends to model Reaches 7 and 8, downstream of Woods Pond to Rising Pond. The exposure concentrations in much of the area downstream of the PSA are lower than in the PSA; therefore, the downstream modeling exercise will explicitly address this issue. The downstream modeling is also advantageous from the perspective of evaluating data sets that are independent of data considered during FCM calibration. Please refer also to the response to General Issue 2.

The temporal capability of the bioaccumulation model is tested in terms of the relationship of the PCB concentration in fish species with age. Other temporal effects (e.g. summer vs. winter) are not explored.

RESPONSE 1-FG-19:

EPA emphasized the age-dependent bioaccumulation of PCBs in fish when evaluating the temporal response of the bioaccumulation model. Nevertheless, seasonal responses in model subcomponents (e.g., temperature-dependent growth and respiration) were considered during model development.

The ability of FCM to track the response of biota to inter-annual changes in PCB exposure concentrations could not be evaluated quantitatively during model calibration because PCB concentrations in exposure media were relatively stable across the 5-year calibration period. The seasonal responses of fish species (e.g., summer versus winter) also could not be explored in detail because there were only a few fish samples collected during winter and spring months. Sampling of fish by both EPA and General Electric has historically emphasized the late summer and early fall in order to control for potential seasonal influences (e.g., transient effects of reproduction/spawning). Some largemouth bass samples were collected in May 1999 to support the fish reproduction study performed for the ERA, but the sample sizes were not sufficient to conduct seasonal comparisons with adequate statistical power.

In addition to considering the seasonal responses to model subcomponents, EPA considered the seasonal patterns in FCM simulation output (i.e., reality checks of seasonal oscillations using professional judgment). EPA also conducted tests of rate parameters (e.g., elimination rates) and gauged the model output against values obtained from the literature. EPA will consider documenting some of these tests in the Model Validation Report.

Overall, though, there appears to be a good PCB concentration data set available to assess the capability of the model. There is a lack of fish tissue concentration data for Reach 5D. However, I do not think that this should preclude the calculation of fish concentration data for fish in this reach. One of the goals of the model is to make estimates where concentration data are not available.

RESPONSE 1-FG-20:

EPA agrees that the ability to simulate concentrations where concentration data are not available is an important goal of the modeling exercise. The uncertainty associated with Reach 5D simulations is discussed in Responses O-DE-7 and 2-JL-1.

Wilbert Lick:

The variations in sediment erosion/deposition between shallow and deep waters of a cross-section of a river can be and generally are quite large. Averaging the suspended solids data and comparing it with the present model which only has one grid cell in it may give a good comparison but will not accurately predict erosion/deposition patterns now and over the long term.

RESPONSE 1-WL-1:

The purpose of the model is not to accurately predict erosion/deposition at a spatial scale smaller than an individual grid cell, and such predictive capability is not necessary to achieve the goals of the modeling study. Please refer also to the response to General Issue 4.

The thickness of the well-mixed sediment layer is an extremely difficult parameter to determine by calibration. Due to its assumed thickness, variations in contaminant concentrations are very slow (see numbers in Question 3) and will be difficult to detect. A determination based on scientific reasoning may be the only way to go.

RESPONSE 1-WL-2:

Please refer to the response to General Issue 6.

E. John List:

My opinion is that the watershed modeling program HSPF is adequately calibrated for use in a statistical sense, although there appears to be insufficient rain gauge information available to apply the model in a storm-by-storm prescriptive way. In other words, the model cannot be used to generate stream flows from a specific storm, but it will be satisfactory when applied to generate synthetic stream flow sequences appropriate for use in the comparison of remediation options. The reason for this is that remediation options will be compared based on a series of stream flow sequences and it is not relevant that these stream flow sequences will ever actually occur, it is enough that they properly represent the statistical variation for the watershed, which it appears they do. The overall strategy of producing a synthetic stream flow record from a rainfall record is a well-founded technique in hydrology and this work appears to have been done well.

However, it is of concern that the HPSF model, apparently because of a lack of adequate rainfall data, was unable to properly represent the two large flow events of May and September in 1999.

RESPONSE 1-JL-1:

EPA agrees that the watershed model is adequately calibrated to produce a representative hydrograph for use in evaluating remedial alternatives. Discrepancies in the representation of the May and September 1999 storms resulted in significant efforts to improve the simulation of those events without adversely impacting the simulation of the other years and events. After numerous efforts to improve (i.e., increase) the simulated flow for those events, it was concluded that any further calibration would lead to parameter values that would not be appropriate for the remaining years of the calibration period.

Examination of the annual flow simulations in Attachment A.4 (Figures 5 through 15 for Coltsville, and Figures 20 through 30 for Great Barrington) indicates that the 1999 simulation is the worst of the 11 year simulations at Coltsville, but one of the better annual simulations at Great Barrington; the percent error values in Table 2-5 also confirm this conclusion. It is likely that the methodology used to assign rainfall above Coltsville resulted in an underestimate of the actual amount of rainfall for 1999. The rainfall patterns and amounts for the other years in the calibration were adequately represented, resulting in a very good overall calibration.

It is not uncommon for this type of situation to occur in watershed modeling, where unusual storm tracks, patterns, and movements in selected years can lead to inadequate, or excessive, estimates of rainfall. This is the primary reason that calibration over many years is recommended for watershed modeling.

On the other hand, the PCB fate and transport model EFDC application appears to have limitations in that the variance of the model output for water column tPCB does not reflect the variance in the measured data. This failing of the model is encapsulated in the comparisons presented in Figure 5-30, where it can be seen that the variance in the measured data is about two orders of magnitude larger than the predicted variance. Part of this problem is most likely associated with the high degree of variance that exists in the sediment PCB concentrations throughout the study area, as illustrated in Figure 5-26.

RESPONSE 1-JL-2:

Please refer to the response to General Issue 10.

Another issue with the EFDC modeling calibration is the relative lack of data that appear to exist in the last nine months of the calibration period. In reality, it is a stretch to claim that the model has been calibrated over a 14 month period when so few data exist for the last nine months, (see Figures 5-17, 5-18, and 5-19), especially since it appears that the intention is apply the model to time scales at least 20-30 times longer than the actual nine month calibration period.

RESPONSE 1-JL-3:

The data richness for the last 9 months of the calibration period is typical for the majority of the timeframe from 1996 to present. Prior to 1996, even fewer data exist to make model-data comparisons. EPA does not agree with the implication by the Reviewer that these data are insufficient to be useful; they are typical of what is available for most modeling studies.

A longer calibration period is clearly appropriate.

RESPONSE 1-JL-4:

Please refer to the response to General Issue 2.

The comparison of average PCB mass flux with predicted mass flux, over the 14-month calibration period that is shown in Figure B.4-46, is essentially meaningless because there is no indication of the confidence limits that apply to either the field estimates of the flux or the predicted flux. On the face of the data presented it appears that the prediction is valid but in the absence of any error margins for both the field and modeling estimates the results in Figure B.4-46 cannot be used legitimately as a calibration basis for the EFDC model.

RESPONSE 1-JL-5:

The average PCB mass fluxes simulated by the model are presented in Figure B.4-46 to convey a general understanding of PCB transport in the Housatonic River PSA. PCB mass fluxes derived from an empirical analysis (the flux analysis) are presented in the figure to provide a qualitative comparison between the simulated concentrations and data. The fact that the two different approaches show similar patterns of PCB transport through the PSA provides support to the conclusion that the model results are reasonable. This summary represents only one of several types of comparisons used in the multiple lines of evidence approach to model calibration.

EPA does not agree that Figure B.4-46 and other presentations that do not include error statistics are "meaningless." The majority of fate and transport models, including EFDC, are deterministic, not probabilistic; therefore, they do not produce predicted parameter values with associated confidence limits. The use of confidence limits to compare model predictions with data, however preferable such comparisons may be, is the exception rather than the norm in numerical fate and transport modeling. It may be possible to add approximate "error margins" to the comparison based on the uncertainty analysis; the uncertainty analysis for EFDC is being conducted as part of the Phase 2 calibration effort. Please refer also to the response to General Issue 11.

A similar disparity in the variance of the predicted and measured PCB tissue concentrations is an outcome of the FCM modeling, as illustrated by Figures C.3-50 and C.3-51. However, the overall performance of the FMC model, as depicted in Figure C.3-51, is certainly encouraging, given the difficulty associated with such food chain modeling. Nevertheless, it would be very nice to have this disparity resolved.

RESPONSE 1-JL-6:

EPA agrees that the central tendencies simulated by FCM are encouraging, as indicated by the achievement of the model performance criteria specified in the MFD/QAPP. The simulated fish tissue concentrations in Figures C.3-50 and C.3-51 are based on central tendencies of model input parameters, not distributions. Therefore, the individual variations in fish bioaccumulation are not represented by the FCM results shown in the figures, and comparisons against field measurements are not appropriate. The model simulations will always have lower variability than the field data. For the Model Validation Report, EPA will explore alternative presentation formats and will revise the description and interpretation of model-data comparisons to improve clarity. Refer to Responses O-DE-9 (above) and 2-FG-4 (below) for additional discussion of the interpretation of "simulated versus measured" PCB scatterplots. Please refer also to the response to General Issue 10.

The basic argument, if I understand it correctly, is that the model is only modeling average fish and not individual fish so the field variance in individual fish cannot be properly represented. I see this as a good reason to adopt a Monte Carlo approach that samples from populations of fish and the creatures that fish feed on. This Monte Carlo method has now become the norm in human health risk analysis and could easily be applied in this context.

RESPONSE 1-JL-7:

Probabilistic analyses were performed as a component of both the human health and ecological risk assessments. However, the PCB concentration in fish (or other) diet was not treated as a distribution but as a point estimate in these documents for two reasons. First, to be consistent with EPA's probabilistic risk assessment guidance, the 95th UCL of the mean was used and, second, because both human and ecological receptors integrate exposures over both space and time, thereby making the average concentration the appropriate concentration of interest. Please refer also to the response to General Issue 1.

EPA agrees that variability in field measurements of individual fish cannot be represented using the deterministic application of FCM. Monte Carlo approaches have some value for quantitatively describing variations in individual fish. However, Monte Carlo model techniques are also limited in the sense that they do not explicitly discriminate between model variability (i.e., observed variations in parameter values observed in the field, such as lipid contents) and model incertitude (i.e., uncertainty regarding the parameters that drive model simulations, such as chemical assimilation efficiency). Furthermore, as described in Comment 4-FG-5, the application of Monte Carlo techniques to complex models with correlated inputs is not straightforward.

The complexity inherent in the application of Monte Carlo techniques is illustrated in the following example, which considers only one sensitive FCM input parameter (organic carbon [OC]-normalized PCB concentrations in sediment). To model the variability in individual largemouth bass, it would be necessary to estimate the distribution of PCB concentrations in sediment to which the base of the food web is exposed. Because largemouth bass are territorial and forage over limited areas, the estimation of exposure sediment concentration distributions for individual fish must consider the variability in sediment concentrations observed among numerous subsections (home ranges) within each river reach. In this example, input distributions were estimated as follows:

- All available surface sediment samples collected in the main channel of the PSA since 1990 were used to construct distributions of OC-normalized tPCB sediment concentrations.
- Home range sizes for largemouth bass were estimated from the literature, and the river was broken up into segments or grids to approximate possible fish home ranges within each study reach.
- All sediment samples within each segment were averaged, and the set of these averages formed the sediment distribution for each study reach.

The resulting distributions for sediment exposures are shown in Figure 1-JL-7. The figure suggests that sediment PCB exposure concentrations are likely to be highly variable among fish occupying different segments of the same PSA reach. This explains in part why PCB concentrations in individual fish specimens can vary by one or two orders of magnitude even within the same species and river reach.

The example illustrates that consideration of the input distributions of sensitive FCM parameters can help to explain the large variability observed in individual fish PCB measurements, relative to central tendencies simulated in the deterministic FCM simulations. However, EPA does not believe that a comprehensive Monte Carlo analysis to simulate distributions of individual fish is warranted, for the following reasons:

- The driver for the evaluation of remediation alternatives is not individual fish, but rather average fish concentrations at spatial scales over which humans and wildlife integrate their fish consumption (see response to General Issue 1).
- The specification of input distributions usually includes a large number of uncertainties. In the above example, the estimated size of bass home ranges, treatment of statistical outliers, treatment of non-detected values, and other data processing considerations introduce uncertainty into the analysis. A number of simplifying assumptions are required for these analyses that can strongly influence the results. Accordingly, it is difficult to discern the

variability for each input parameter from the incertitude inherent in the specification of each distribution.

- The incertitude described above is compounded by the number of input parameters in FCM, a number of which are intercorrelated. Accounting for multiple intercorrelated inputs further increases the complexity of the analysis and limits the degree to which Monte Carlo simulations can be quantitatively compared against field data distributions.

In summary, EPA does not believe that incorporation of a probabilistic model around the FCM simulations to predict concentrations in individual organisms is necessary, given the considerable effort and complexity of the task, the high uncertainty associated with the output, and the goal of the modeling study. The variability in individual field measurements is not unexpected given the large number of variables that mediate the exposure and bioavailability of PCBs to fish.

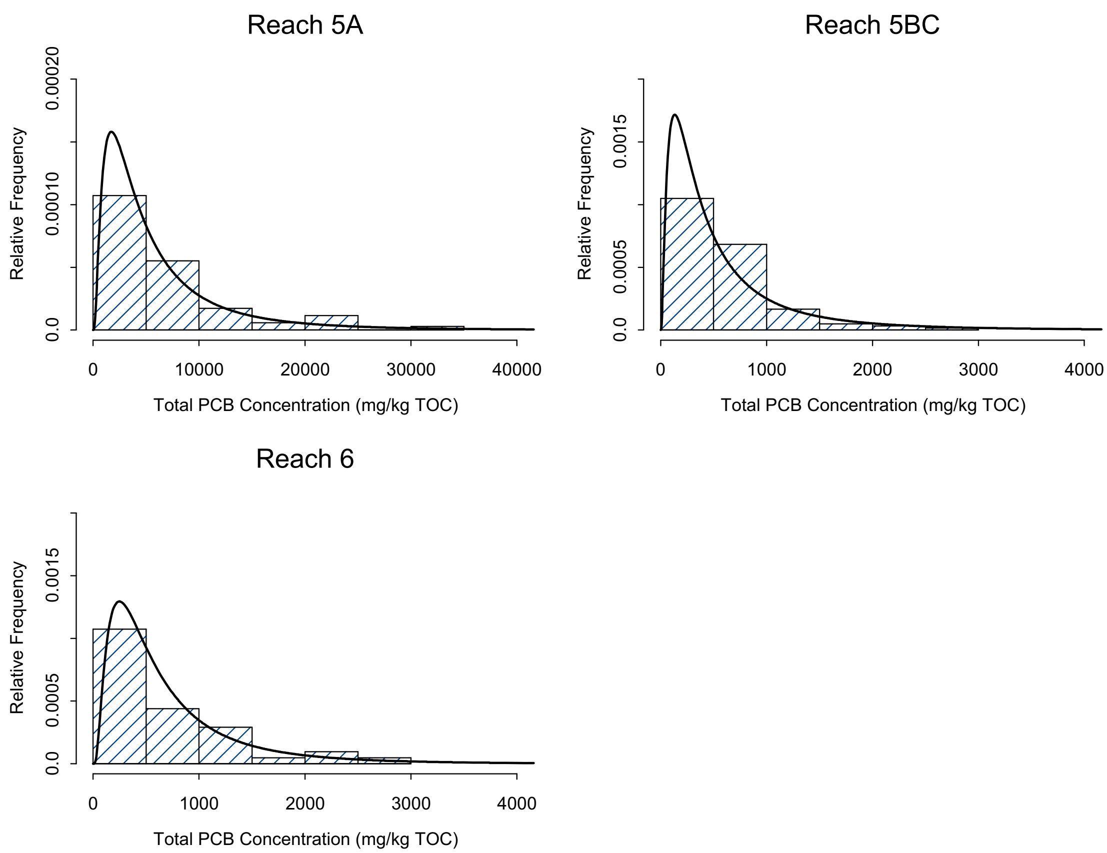


Figure 1-JL-7 Histograms of Estimated Average TOC-Normalized PCB Concentrations for Largemouth Bass Home Ranges

Evaluation for Evidence of Bias

2. Is there evidence of bias in the model, as indicated by the distribution of residuals as a function of the independent variables?

E. Adams:

Comparison of predicted and measured mean values is reasonably good for most variables, as would be expected following calibration. But, as the discussion above implies, the models can be expected to under predict the fluxes of sediment and PCBs under extreme events, and these extremes will be most responsible for changes in PCB concentrations in the future. This is a type of bias.

RESPONSE 2-EA-1:

EPA agrees with the Reviewer's comment that good agreement has been achieved between mean model simulations and data; however, EPA disagrees that strong agreement between predicted and measured mean values is automatically "expected" for any model calibration. This would only be the case if the model parameters are deliberately tuned during calibration to maximize the fit (i.e., minimize the model residuals). The majority of the parameters in the modeling study were specified either as fixed values (a priori, based on site data or literature) or were only allowed to vary within limited and scientifically plausible ranges.

EPA disagrees with the Reviewer's expectation that simulated fluxes of sediment and PCBs will be underpredicted for extreme events. The Reviewer's comment, "But, as the discussion above implies …" refers to an inaccurate assumption about the approach used to calculate PCB concentrations of each of the size classes. Response 1-EA-15 discusses how the modeling approach recognizes and accounts for the preferential sorption of PCBs to finer sediment and uses formulations for erosion that account for bed composition. These approaches avoid the potential problem discussed by the Reviewer.

W. Frank Bohlen:

The distribution of residuals provides no indication of model bias.

RESPONSE 2-FB-1:

EPA agrees with the Reviewer's assessment.

Douglas Endicott:

The models appear to be fairly unbiased in terms of the principal state variables. The following exceptions were noted during review:

- The gradient in total PCB concentrations across Woods Pond indicates bias, due possibly to the magnitude of the diffusive flux from sediment. The net loss of PCBs from Woods Pond contradicts expectation and conceptual model;

RESPONSE 2-DE-1:

The good agreement between the longitudinal profiles of model results and data collected at the lowest flows (Figure B.4-37) indicates that the representation of the diffusive flux in the model is reasonable because any problem with the diffusive flux would be highlighted during the low-flow conditions when dilution is reduced. However, EPA agrees that the model does underpredict the amount of settling that occurs across Woods Pond during a somewhat higher-flow condition (Figure B.4-38). Therefore, further effort will be directed toward refining the particle settling rate formulations used in the TSS model over the course of the Phase 2 Calibration. This effort will include consideration of the flocculation and settling model that has been recommended by Dr. Lick. Calculation of additional solids deposition in Woods Ponds would result in additional accumulation of PCBs.

- Dissolved PCB concentrations are consistently overpredicted in the storm event periods, suggesting either a weakness in the partitioning model or inadequacy in estimating low-end censored concentration data;

RESPONSE 2-DE-2:

EPA agrees in general with the Reviewer's observations, but does not agree that this indicates a weakness in the partitioning model. Most of the analytical results for dissolved PCBs collected during the storm events were non-detect, with detection limits near 0.01 μg/L. During the smaller storm events, the simulated dissolved PCB concentrations were higher than the detection limit of samples reported as non-detect by approximately a factor of two to four. For the data collected at higher flow conditions, simulated dissolved concentrations were typically within a factor of two of the data or detection limit. Because of the relative lack of importance of the surface water dissolved exposure route in the bioaccumulation calculations for this system, the comparisons between simulated dissolved PCB concentrations and data were considered acceptable.

Given the overall ability of the three-phase partitioning model to reproduce the surface water partitioning data set, EPA believes that the partitioning model is adequate.

- There also appears to be an unexplained factor in the sediment pore water partitioning data, possibly some kind of solids effect. Cross-plots of particulate organic carbon versus apparent

K_{oc} show that K_{oc} declines with increasing f_{oc} for both total PCB and congeners, regardless of whether data from some sediment cores are censored. Dissolved and particulate PCB predictions look OK in comparison to the partitioning study data, but this is a very limited number of measurements.

RESPONSE 2-DE-3:

The effect noted by the Reviewer is evident in some of the sediment data, primarily the approximately 15% of the partitioning samples with low f_{oc}. This effect may reflect the relative increase in importance of inorganic surfaces as adsorption sites for PCBs in sediment with low organic carbon content, and the fact that this fraction of sorbed PCBs is included within the fraction sorbed to organic carbon in the 3-phase model.

This concept is demonstrated by considering a limiting condition where the carbon-normalized PCB concentration approaches infinity as the organic carbon content approaches 0, even though the total dry weight PCB concentration and the pore water concentration may be very low in such samples. The result would be an apparently high carbon-normalized PCB concentration and partition coefficient. The addition of an inorganic phase to the 3-phase representation of PCB sorption was considered; however, it did not provide additional predictive capability. The samples collected for the partitioning data set were intentionally biased to obtain a disproportionate number of low f_{oc} cores. On a mass basis, far more PCB mass is found in high f_{oc} areas (where carbon normalization works well) than in low f_{oc} areas; therefore, a decision was made to retain the 3-phase representation of PCB sorption. Please refer also to the response to General Issue 5.

- Some mild bias is evident in predictions of PCB in fish. Predicted PCB concentrations in bullhead, sucker, sunfish and bass are generally lower than mean observations, while cyprinid PCB concentrations are overpredicted in all reaches except 5A. There is really too little data to check bias in lower food chain predictions.

RESPONSE 2-DE-4:

Statistics that represent bias (e.g., mean error and mean percent error in Table C.3-16 and the MB* statistic derived in Response 2-FG-5) do not support the Reviewer's comments. The referenced fish species are discussed below, in relation to the statistical and graphical measures of bias presented in Appendix C:

- Brown bullhead – Figure C.3-31, Figure C.3-34, and Figure C.3-36 all indicate that simulated and measured PCB concentrations are very similar, on both a wet weight and lipid-normalized basis. The mean error is slightly negative (-6.2 mg/kg) and the MB* statistic is only slightly above one (1.28), both of which indicate minor *overprediction* of observed concentrations. From this information it is clear that simulated concentrations are not "generally lower than mean observations" as suggested by the Reviewer; the reverse may be true depending on the definition of "mild."

- Sunfish – Figure C.3-31, Figure C.3-34, and Figure C.3-37 all indicate that simulated and measured PCB concentrations are very similar. In Figure C.3-31, tPCB concentrations in Reach 6 (Woods Pond) sunfish are slightly overpredicted, concentrations in Reach 5B/5C sunfish are slightly underpredicted, and measured and simulated values in Reach 5A sunfish are nearly identical. The mean error for pumpkinseed sunfish is slightly negative (-5.5 mg/kg) and the MB* statistic is only slightly above one (1.30), both of which indicate minor *overprediction* of observed concentrations. From this information, the Reviewer's conclusion that tPCB concentrations in sunfish are "generally lower than mean observations" is not supported by the data.

- Largemouth bass – As shown in Figure C.3-31, the measured concentrations in bass are slightly underpredicted in Reaches 5A and 5B/5C, and are slightly overpredicted in Reach 6 (Woods Pond). The mean percent error of -21% and the MB* statistic (1.28) both indicate a small overall *overprediction*. From this information, the Reviewer's conclusion of mild bias toward underprediction is not supported by the data.

- White sucker – As shown in Figure C.3-31, the average simulated concentrations in all reaches are slightly lower than the measured concentrations. Consequently, the MB* statistic is slightly below one (0.82) and the mean percent error is positive (+20%), indicative of slight underprediction. EPA agrees that this pattern could be interpreted as a "mild bias." Although this bias could be removed by conducting further calibration of the model, this was not done because the model results were already well within acceptable tolerance levels and because it was believed that "tuning" of the model in this manner would not improve the predictive value of the model.

- Cyprinids – EPA agrees that cyprinids were systematically overpredicted in reaches downstream of Reach 5A, based on the 10 golden shiner tissue samples. The overprediction of downstream cyprinids is acknowledged on page 6-31 of the Model Calibration Report, which states:

  > "The FCM results for cyprinids are higher than the field measurements of golden shiners made in downstream PSA reaches. Golden shiners have a diet dominated by water-column-based prey; therefore, the mixed diet used in FCM for the cyprinid community is not intended to apply to golden shiners. Overall, FCM is not sensitive to this overprediction because the contribution of cyprinids to the diet of predators is minimal in downstream reaches."

With respect to the "lower food chain predictions," EPA agrees that the sample size of D-net invertebrate measurements is low relative to fish samples. However, the FCM predictions of tPCBs in epifauna (mg/kg ww; depicted in Figure C.3-9) were all close to the PSA mean invertebrate concentrations in

D-net samples, which consisted primarily of epifauna. Furthermore, the FCM simulations were also consistent with the crayfish and tree swallow diet concentration data presented in Section C.3.4.2. EPA believes there are sufficient data to conclude that there is no significant bias for epifaunal invertebrates.

I would not judge any of these biases to be so significant as to undermine the credibility of the models.

RESPONSE 2-DE-5:

EPA agrees with the Reviewer's assessment.

Marcelo H. Garcia:

The field observations indicate high variability of PCB concentrations over very small spatial scales. The model is incapable of capturing such variability. Thus if the model has a bias, it is toward the "filtering" of the high variability seen in the field.

RESPONSE 2-MG-1:

EPA agrees that the models will never reproduce the amount of variability observed in the field. Please refer to the response to General Issue 10.

At the same time, it is not clear how much of the observed variability is indeed natural and/or the result of the sampling protocol.

RESPONSE 2-MG-2:

Based on the rigorous sampling protocols and extensive QA implemented for the Rest of River study, EPA does not believe that any portion of the observed variability is an artifact of the sampling protocol. However, as discussed in the RFI Report (BBL and QEA, 2003) EPA previously identified a bias (low) in the sediment PCB data analyzed by the on-site laboratory, believed due to analytical protocols for sample extraction. Please refer also to the response to General Issue 5.

What is clear is that the model will not capture such "randomness" and then the question is if this capability is indeed necessary for the intended use of the model.

RESPONSE 2-MG-3:

EPA agrees that the model will not reproduce the variability observed in the data, and does not believe that it is necessary to do so to achieve the objectives of the modeling study. Please refer to Response 1-EA-11 and the response to General Issue 1.

Frank Gobas:

Figure 3-18 shows no bias in the distribution of residual flows as a function of the measured flow rates.

RESPONSE 2-FG-1:

EPA agrees with the Reviewer's assessment.

Figure 4.36 also shows no significant systematic bias for the calculations of TSS by the HSPF model although some considerable variation between observed and predicted TSS values was found in some cases.

RESPONSE 2-FG-2:

Figure 4-36 refers to residual and relative residual TSS simulated by EFDC, not HSPF. EPA agrees with the Reviewer that the residuals and relative residuals depicted in the figure indicate no systematic bias in the model results. EPA acknowledges that there are some residuals that indicate occasional discrepancies between model results and the data, but notes that the largest discrepancies are for relative residuals, which are a less important indicator of the utility of the model than actual residuals. EPA believes that the magnitude of the differences is not unusual for this type of modeling study, and does not indicate an issue with the overall adequacy of the model calibration.

The model necessarily simulates a TSS concentration that is integrated over space (i.e., over at least the area of a single grid cell) and over time. A measured datum reflects the concentration at a much smaller spatial scale and at a very discrete time. Accordingly, discrepancies between a simulated concentration and a measured concentration are to be expected. Of greater importance is whether the central tendency of the simulation is accurately depicting the central tendency of the data. As shown in Model Calibration Report Figures 4-21 through 4-24, and in other similar figures in Appendix B, EPA believes this to be the case.

With regard to the magnitude of the deviations, it is important to keep in mind that some of the events were sampled intensively over time, at selected stations. As a result, a slight discrepancy in the timing of an event could lead to marked deviations in point-wise comparisons of simulated and measured TSS levels, even though the pattern of the response over the course of the event is relatively well represented (e.g., Event 6, Figure B.3-50, at New Lenox Road). Many of the variations referred to by the Reviewer are related to results from storm events affected by timing rather than the magnitude of the response.

Estimates in the tPCB water column concentrations show no significant systematic bias at Holmes Road and Woods Pond Headwaters but some bias is apparent from the distribution of the residuals for data collected at New Lennox Road and Woods Pond Outlet (Figure 5-31).

RESPONSE 2-FG-3:

The slight indication of bias in the results is only evident at the New Lenox Road and Woods Pond Outlet locations, and only with respect to the relative residuals. As discussed in Response 2-FG-2, above, EPA believes that actual residuals are a more important measure of model performance and that the indication of slight bias is of limited significance for achieving the overall goals of the modeling study.

As was the case for TSS, it is important to recognize that the event data were sampled relatively intensively for a few of the events, at selected stations. For example, in the case of New Lenox Road, numerous samples were analyzed for PCBs during Event 1 (May 19-21, 1999, Figure B.4-44). The concentrations during the rising part of the hydrograph for this event were underpredicted due to a slight phase shift, and peak concentrations were also somewhat under-predicted. This phase shift leads to differences in point-by-point comparisons of model and data, even though the pattern over the course of the event is judged to be well-represented by the model results. A slight error in timing can lead to a large deviation of the model from the instantaneous concentration data, while having an inconsequential impact on both the mass loading of PCBs from the PSA and the response of the food chain to the transient change in concentration.

Figures C.3-26, C.3-27 and C.3-29, and C.3-49, C.3-50 and C.3-52 (for the linked model) plot the residuals against the measured PCB concentrations. The plots do not lend themselves to explore issues of bias. A statistical treatment of the data would be more useful. Hence, it is difficult to confirm the statement on p. C.3-31 that there is no model bias across the range of PCB concentrations evaluated. Just looking at Fig C.3-26, it looks as if there are more data points below the zero line than above it. Figure C.3-27 appears to confirm this for coplanar PCBs. But again, this may not be so.

RESPONSE 2-FG-4:

EPA agrees that the figures cited above do not, in isolation, allow for a comprehensive evaluation of model bias. However, other materials supplied in Appendix C allow for a more rigorous evaluation of model bias. The following information should be used in conjunction with these figures in evaluating bias:

- Figures C.3-28 and C.3-51 are more appropriate graphics for assessment of bias because they compare central tendencies of field data (shown in yellow square symbols) to central tendencies simulated by the model.

- Attachments C.15 and C.16 provide numerous figures of simulated versus measured concentrations, as a function of age, reach, and species. These plots, expressed in both wet weight and lipid-normalized units, provide considerable information on the overall model performance and bias.

- The quantitative measures of model performance presented in Section C.3.6.4 (Tables C.3-14 through C.3-16) provide the "statistical treatment" of data requested by the Reviewer.

EPA believes that the model performance measures presented in the report (both graphical and quantitative), provide sufficient information for assessment of model bias. However, in Response 2-FG-5 (below) EPA has provided supplemental quantitative information.

There are various ways to explore the issue of bias for PCBs on a congener or total-PCB basis. We have used the model bias MB, which is derived on a species-specific basis as:

$$MB_j = 10^{\left(\sum_{i=1}^{n} \frac{\left[\log\left(C_{P,i} / C_{O,i} \right) \right]}{n} \right)}$$

In essence, MB_j is the geometric mean (assuming a log- normal distribution of the ratio $C_{P,i} / C_{O,i}$) of the ratio of predicted ($C_{P,i}$) and observed ($C_{O,i}$) for all PCB congeners i in a particular species j included in the analysis. MB is a measure of the systematic over- (MB>1) or under-prediction (MB<1) of the model. It should be stressed that in the calculation of MB, over- and under-estimations of the observed concentration values for individual PCB congeners have a tendency to cancel out. Hence, MB tracks the central tendency of the ability of the model to predict PCB congener concentrations. It is a useful measure of model performance if total PCBs (SPCB) are of primary interest. The variability of over- and under-estimation of measured values can be represented by the 95% confidence interval of MB (i.e. 95% CI = antilog(geometric mean $\pm$ ($t_{v,\ 0.05} \times$ standard deviation)). The 95% confidence interval represents the range of concentrations that includes 95% of the observed concentrations. It can be viewed as a measure of the uncertainty of the model predictions. The same approach can also be applied to total PCB as well. In that case, model bias MB* is:

$$MB_j^* = 10^{\left(\sum_{i=1}^{m} \frac{\left[\log\left(C_{P,\Sigma PCB} / C_{O,\Sigma PCB} \right) \right]}{m} \right)}$$

MB_{j^*} is the geometric mean (assuming a log- normal distribution of the ratio $C_{P,\Sigma PCB} / C_{O,\Sigma PCB}$) of the ratio of predicted and observed concentrations for ΣPCB in species j (Arnot and Gobas, Environ. Toxicol. Chem. 23, 2343-2355 (2004)).

RESPONSE 2-FG-5:

In response to the Reviewer's suggestions, EPA has reevaluated the model performance using the MB and MB* statistics defined above (using the linked version of FCM) as shown in Table 2-FG-5 below.

Table 2-FG-5

Model Bias Statistics, by Species, for FCM Linked Model

Species	n	MB	n*	MB*
Brown Bullhead	405	1.385	45	1.278
Cyprinid Fallfish	45	0.881	5	0.809
Cyprinid, Shiner	90	2.283	10	2.665
Pumpkinseed	540	1.300	60	1.270
Largemouth Bass	630	1.282	70	1.277
White Sucker	486	0.936	54	0.815

EPA notes that the results of the MB and MB* calculations result in conclusions that are very similar to the mean percent error calculations presented in Tables C.3-14 through C.3-16 of the Model Calibration Report.

The MB* statistic (i.e., model bias for tPCBs) is just over 1.0 for three of six species (bullhead, sunfish, and bass), indicating a slight overprediction and just under 1.0 for two of six species (fallfish, sucker) indicating a slight underprediction. The magnitude of bias for these species is low (i.e., within MFD/QAPP performance targets, and well within the range of error expected for most bioaccumulation models). EPA believes that further adjustment of the models to push the MB* statistic closer to 1.0 would constitute overcalibration and would not necessarily improve model predictive ability. The bias toward overprediction in the remaining fish group (downstream cyprinids) is discussed in Response 2-DE-4.

I do agree with the authors that the vast majority of the observed concentrations are within an order of magnitude of the mean simulated by the FCM. However, it is not clear why Figure C.3-29 only suggests a range of an order of 2 rather than 10 (in Fig C.3-26). The latter may be due to the fact that Fig C.3-26 refers to the combined data set while Fig C.3-29 refers to means of subsets of samples. (This needs to be clarified in the Figure legend).

RESPONSE 2-FG-6:

The error statistics presented in Figures C.3-26 and C.3-29 are computationally different, and cannot be compared in the manner implied by the Reviewer. The equation for mean percent error (as presented in Figure C.3-29) is defined on page C.3-34 as:

$$MPE \quad \frac{100}{n}\sum_{i-1}^{n}\frac{e}{\left(y_i + \hat{y}_i\right)/2} \qquad \text{(C.3-2)}$$

where:

MPE = Mean percent error, a unitless measure of model bias (%).

yi, = Measured concentration.

$\hat{y}_i$, = Simulated concentration from model.

e = Residual (measured – simulated value).

n = Number of model/data pairs evaluated.

In contrast, the equation for the residual of the log-log transformed model (as presented in Figure C.3-26) is defined as:

$$\log(y_i) - \log(\hat{y}_i)$$

These two error statistics result in different quantitative values. For example, for an observed value of 0.1 and a model simulation of 1.0, the derived MPE would be -164%, whereas the residual of the log-log transformed model would be -1.0 (log mg/kg OC). Both statistics are indicative of a “factor of ten” difference between measured and simulated values.

Figure C.3-29 does not indicate a range “on the order of 2,” as suggested by the Reviewer. The graph indicates: (1) that individual PCB concentrations can range up to an order of magnitude or more from the mean model-predicted PCB concentrations; and (2) that there are approximately as many positive as negative model residuals. Because Figure C.3-29 depicts individual data points relative to simulated central tendencies, the magnitude of individual errors does not represent only model error, but also incorporates variability in the data.

While the residuals may not show a bias, Fig C.3-27 illustrates that the uncertainty in the model predictions can be large. It is possible that this uncertainty is to some degree caused by difficulties in modeling the bioaccumulation behavior of certain congeners. Metabolism can be a significant process for some congeners and not for others. It is therefore instructive to explore the issue of bias on a congener specific basis rather than combining many congeners in one analysis. I recommend this is done.

RESPONSE 2-FG-7:

Figure C.3-27 does not depict uncertainty in model predictions. Figure C.3-27, as well as Figures C.3-26 and C.3-29, depict individual data points relative to modeled central tendencies. As a result, the wide range of scatter in the graphs is attributable mainly to the variability in individual measured fish PCB concentrations, not to model uncertainty. Figure C.3-28 demonstrates that when field data are aggregated and represented by a measure of central tendency (to make them comparable to deterministic model simulations), the measured values are generally close to FCM simulations.

EPA agrees that difficulties in the modeling of bioaccumulation of certain congeners can increase the uncertainty of the model. The congener-specific modeling yielded more uncertain results than the tPCB modeling, in part due to uncertainty in congener-specific bioavailability of coplanar congeners, and in part due to a lack of congener-specific information for some model parameters (i.e., estimation based on physical/chemical properties). As a result of this uncertainty, the spread of the data in Figures C.3-28 and C.3-51 is somewhat wider for congeners than for tPCBs.

EPA agrees that, in addition to analyses with all congeners combined, model bias should be considered on a congener-specific basis. Appendix C already contains this information in both graphical and statistical forms. For example, Figures C.3-40 through C.3-48 provide summaries of simulated versus measured congener concentrations in fish, organized by river reach and species. Model fit statistics are also broken down by congener in Tables C.3-8 and C.3-14. From these analyses it was concluded (Section C.3.6.3.2) that most congeners were predicted without significant bias; some exhibited small systematic overpredictions, and some exhibited small systematic underpredictions.

Also, I recommend that, in addition to the mean model bias, the uncertainty around the mean model bias is explicitly stated. A high degree of uncertainty of the model calculations should not be viewed as criticism of the model but a reflection of the actual state of the modeling capability.

RESPONSE 2-FG-8:

EPA will consider characterizing the uncertainty around the model mean bias in the Phase 2 Calibration and Validation reporting using the confidence intervals about the MB statistic, as described in Response 2-FG-5.

Wilbert Lick:

[No comments]

E. John List:

There certainly appears to be bias in the FCM as indicated by the lipid-normalized PCB concentrations predicted by the FCM in Reach 5D of the PSA, as indicated by Figure 19 of Attachment C.16, and other portrayals of the FCM predictions for Reach 5D that are included in Attachment C.16. It is unfortunate that there appears to be no field data to substantiate the predictions of the FCM in this reach of the river, especially since the predictions appear to be incongruent with those for the rest of the PSA. In the absence of any explanation and/or correction for this phenomenon it is not clear that it is real, in which case it makes the application of the model problematic, at least until the reason for this indicated bias is understood.

RESPONSE 2-JL-1:

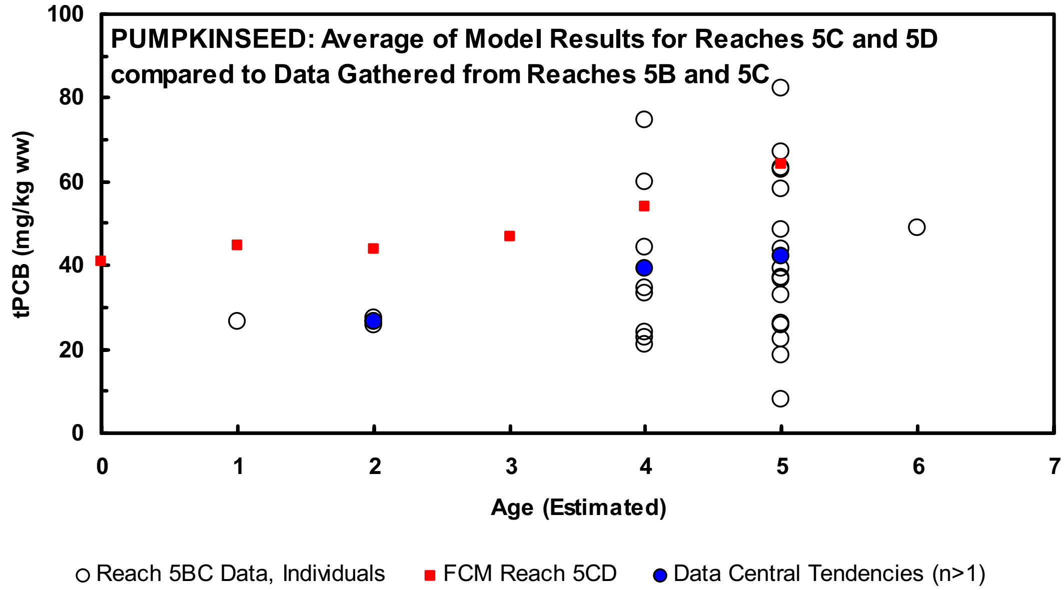
The Reviewer is correct that the highest fish PCB concentrations were simulated for Reach 5D using the linked version of the bioaccumulation model. However, it is incorrect to characterize this observation as "bias" because there are no field tissue data for Reach 5D. Assessment of model bias can be made only if both measured and simulated values are available for comparison.

It appears that the Reviewer is commenting on the discrepancy between Reach 5D simulations and the simulations in adjacent downstream reaches (i.e., Reach 5C and Woods Pond). Although this difference is not bias per se, discussion of the uncertainty associated with the difference, and the implications for model application, is appropriate.

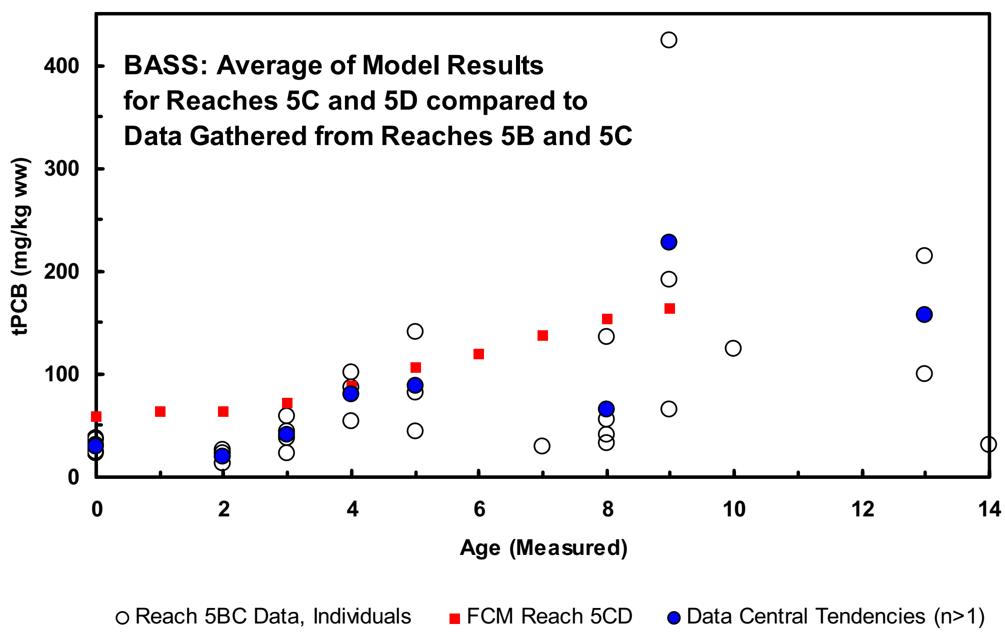
Fish samples were not collected within Reach 5D (i.e., large backwaters contiguous to Reach 5C) because the backwaters were not accessible at the time the samples were collected. However, evaluation of largemouth bass life history indicates that fish migrate back and forth between the backwaters and the main channel of the Housatonic River. Therefore, the fish tissue data collected for Reach 5C represent mixed exposures to contaminated media in both Reaches 5C and 5D, with the degree of backwater habitat use dependent on flow, temperature, dissolved oxygen, and other habitat factors. When the habitat uses of Reaches 5C and 5D are considered and weighted by expected exposure duration, the model predictions remain close to the observed main channel concentrations, and within the tolerance limits for calibration specified in the MFD/QAPP.

Figure 2-JL-1 shows the results of additional linked model simulations in which model estimates from Reaches 5C and 5D were averaged, assuming 50% exposure from each reach. Even after the influence of backwaters exposure is taken into account, the measured and simulated concentrations remain in reasonable agreement.

(a) Pumpkinseed



(b) Largemouth Bass



Notes: Blue symbols represent means of all individual fish for each age class, and are directly comparable to FCM simulations.

Red symbols represent FCM simulations of mean concentrations for each age class.

Open symbols represent individual measured PCB concentrations; these data are portrayed to illustrate the variability and sample size of data for each age class.

Figure 2-JL-1 Comparison of Measured Fish PCB Concentrations to FCM Simulations Assuming Mixed Habitat Use of the Mainstem and Associated Backwaters

Ability of the Model to Account for Relevant Processes

3. Does the model, as calibrated, based upon your technical judgment, adequately account for the relevant processes affecting PCB fate, transport and bioaccumulation in the Housatonic River?

E. Adams:

All three models are considered state of the art and hence come generally "fully equipped". The model team is experienced and has added important features such as the effects of vegetation on stream flow and transport. Hence I believe that most processes are at least *represented* in the models. I do agree with QEA/GE that bank erosion/river meandering is an important process that should be included and could help explain the large vertical spread of PCBs observed in cores.

RESPONSE 3-EA-1:

Please refer to the response to General Issue 9.

Having said that, it is not clear that the models have been properly calibrated when it comes to individual processes. While each of the three models has their own unique calibration issues, EFDC is perhaps the most problematic because: 1) it is relatively new and has not been used in the current framework (e.g., with both in-channel flows and above bank flows),

RESPONSE 3-EA-2:

Although it is true that some capabilities of EFDC being used in the application for the Housatonic River are relatively new (e.g., bed load transport), the version of EFDC that is being applied to the Housatonic River has undergone extensive QA by the EPA National Exposure Research Laboratory in Athens, GA, and the model application team. In addition, EFDC is being applied at other contaminated sediment sites where both in-channel flows and out-of-bank flows are being simulated, e.g., the Kalamazoo River, Michigan. The modeling team is not aware of any other contaminant fate and transport model that has been used previously in a framework similar to the Housatonic River (including simulation of out-of-bank transport).

2) it is being calibrated over a short period of time (14 months) relative to the time constants of some of the biochemical processes,

RESPONSE 3-EA-3:

Please refer to the response to General Issue 2.

and 3) compromises are being made because of computational expense.

RESPONSE 3-EA-4:

The Conceptual Site Model for the PSA (WESTON, 2004) includes a number of processes, such as PCB transport onto the floodplain, that require a computationally demanding modeling framework. Accordingly, the fate and transport model selected for the modeling study is computationally demanding. The use of any such computationally intensive model necessarily requires compromise; EPA believes that such compromises have been made in a manner that is consistent with achieving the goals of the modeling study.

Including data from the earlier years (~1980-1999), as has been mentioned for the next phase, will help with the calibration. If this is done it would be nice if the last few years (2000-2004) could be set aside for a true validation (no more parameter tweaking).

RESPONSE 3-EA-5:

EPA has revised the approach to model calibration and validation, in part to address this Reviewer's comment; please refer to the response to General Issue 2.

One issue with calibration is that multiple variables contribute to a particular observation so the net effect may be correct, while the individual effect is not. As mentioned previously, one example is water column TSS concentrations that reflect both resuspension and deposition.

RESPONSE 3-EA-6:

Please refer to the response to General Issue 7.

I am still a bit uncomfortable about the fact that most of the PCB mass is in the floodplain (and hence affected only by relatively rare flood flows), yet the vast majority of the computational time is taken with in-channel flows. Should the same model be used (in the same way) for both?

RESPONSE 3-EA-7:

The Reviewer is correct that the majority of the PCB mass resides in the floodplain. However, most of the PCB mass that is subject to scour during a flood event is in the river channel. Most of the computational time is associated with simulation of conditions in the river channel, mainly because the river is within its banks for the majority of the time and when the floodplain cells are dry computations in the cells are not performed. The river sediment serves as a source of PCBs to the floodplain, whereas the PCBs in the floodplain are relatively stable in comparison.

It is not clear whether storms or the routine flows will be most responsible for PCB transport, but the fact that the model can only afford one grid cell over the channel width and a couple of grid cells in the vertical seems to defeat the purpose of a 3D model.

RESPONSE 3-EA-8:

EPA has not characterized this model application as being three-dimensional. The question of the optimal dimensionality for the hydrodynamic/sediment-contaminant transport and fate model was discussed extensively in Section 16.4.2 of the MFD Responsiveness Summary (WESTON, 2002). In that document, EPA restated its belief that a two-dimensional application was the most desirable approach, but indicated that the selected model would be tested in both two-dimensional and three-dimensional modes and the results presented in the final MFD. The results of the testing were presented in Section 5.4 of the final MFD, and the final computation grid was specified as two-dimensional (see MFD, p. 5-56). Please refer also to the response to General Issue 4.

Reference:

WESTON (Weston Solutions, Inc.). 2002. *Responsiveness Summary to the Peer Review of the Modeling Framework Design and Quality Assurance Project Plan: Modeling Study of PCB Contamination in the Housatonic River.* Prepared for U.S. Army Corps of Engineers and U.S. Environmental Protection Agency.

The mass transfer coefficient used to compute sediment-water flux is being calibrated to match observed changes in contaminant flux between two stations. Since the model has only one grid cell per river width, the calculated flux is based on a cross-sectional average flow rate, and hence an average velocity, bottom shear stress and erosion potential. It is quite possible that at a given time portions of a reach are eroding while others are depositing and it is not clear if a calculation based on average flow is correct. Furthermore, I would expect to have seen the mass transfer coefficient increase with river flow, but this apparently was not the case.

RESPONSE 3-EA-9:

Please refer to the responses to General Issues 4 and 8.

I am also concerned about the bioturbation coefficients. I realize that limited actual calibration has taken place so far, but I wonder if true calibration will be possible in the future. The tentative value of E-9 m2/s is quite large; are there biological observations to support such a value (or any value)? 14 months is too short a period to determine if bioturbation (plus diffusive flux across the interface) will have much effect; will adding an additional 10-20 years be that much more helpful, especially if the coefficient turns out to me much smaller?

RESPONSE 3-EA-10:

Observations regarding bioturbation in the literature have been reviewed for applicability to the organisms observed in the Housatonic River, and the information is summarized in the response to General Issue 6. These data will be used to parameterize the long-term calibration and validation of the model (i.e., over simulation periods longer than 14 months). These analyses will, in combination, provide improved support for the evaluation of the particle mixing rate(s) assigned in the model.

Finally, the measured PCB sediment concentrations are not an ideal data set for calibration. As mentioned above, one problem is that there is not much change over (14 months of) time. There is some change over space, but this is overwhelmed by the much greater change over very short (sub-grid scale) lengths, which cannot be resolved.

RESPONSE 3-EA-11:

Please refer to the responses to General Issues 2 and 5.

Finally, under historical conditions PCBs have been ubiquitous, appearing with the upstream inflow, and eroding/diffusing off the sediment bed, river banks, and floodplains. These PCBs "all look the same" and hence it is hard to diagnose transport mechanisms based on a model's ability to "match" them. It was not discussed in detail at the Peer Review Meeting, but if there is time, serious consideration should be given to a tracer experiment of some sort. The best type of experiment, in theory, would be one in which sediment of different types (or placed in different locations) were uniquely labeled (e.g., with fluorescent colors) so that they could be tracked. This would take time and resources to think through, but it may be the only way to calibrate certain model parameters.

RESPONSE 3-EA-12:

Dye studies and the use of other tracers were considered by EPA earlier during the modeling study, and rejected because of lack of public acceptance.

An alternative approach would be to conduct a trial remediation on a patch of river and see how rapidly contaminated sediment from upstream fills in the clean spots. (In principle one could simulate the effect of clean upstream sediments that will result from ongoing remediation to see how fast they fill in the contaminated portions downstream, but it is better to simulate a clean downstream spot since a bit of contamination in an otherwise clean patch shows up more than a bit of clean sediment in an otherwise dirty patch, especially given the variability.)

RESPONSE 3-EA-13:

EPA will consider including a simulation of a hypothetical remedial scenario in the Model Validation Report to illustrate model performance. Please refer to the response to General Issue 12.

W. Frank Bohlen:

As noted, **The Model** consists of three primary components, a watershed model (HSPF), a hydrodynamic-sediment/contaminant transport model (EFDC), and a bioaccumulation model (FCM). This combination of models accounts for all relevant processes affecting PCB transport and fate in the Housatonic River study area. That is, all relevant processes can be accommodated in the models and have received some consideration in model development. Unfortunately this does not mean that this combination of models, as presently structured, will provide the accurate simulations of PCB transport and fate needed to facilitate remedial designs.

RESPONSE 3-FB-1:

EPA recognizes that model code selection alone does not necessarily ensure that the final calibrated model will be adequate to achieve the goals of the modeling study. The objective of model selection was to provide a framework for the modeling that, if properly implemented, would include the relevant processes and achieve the goals of the modeling study.

Determining which processes in the PSA are relevant for contaminant transport and fate, and then selecting model code that includes all relevant processes was only one aspect of model selection. Of equal importance was the question of how the processes are simulated numerically within the model code, i.e., the model formulations. Both of these aspects of model selection, along with other considerations, were carefully evaluated in the selection of the three models used in the Housatonic River modeling study. That process is documented in Sections 4 and 5 of the Modeling Framework Design. EPA believes that the model framework that was selected for this study will achieve the overall modeling study objective of providing a predictive tool for use in the evaluation of remedial alternatives.

The watershed model (HSPF) has the potential to provide both estimation of surface water volume inflows as well as water temperature and the associated sediment loads entering the study area. The development of the model and subsequent comparisons of model outputs vs. measured flows indicated close agreement, well within the QAPP specified "very good" category of +/-10%. The agreement for the case of the suspended solids load was poorer which is not entirely unexpected. Unfortunately the discussion of the reasons for these latter differences was weak.

RESPONSE 3-FB-2:

Please refer to Response 1-FG-3, above.

Moreover, the comparisons relied on relatively long term averages (mean annual load). Examination of higher frequency TSS data/model comparisons for several storm events (May and September, 1999 e.g.) show differences exceeding 100% in concentrations as well as significant differences in timing. These differences make one wonder how it is that the annual average data is so well simulated. This was not subject to sufficient discussion in either Volume 1 or Appendix A.

RESPONSE 3-FB-3:

It is not reasonable to expect a watershed model to predict suspended solids on the time scale of individual events. As stated on page 2-38 of the Model Calibration Report, the watershed model TSS calibration was based on comparisons to the annual loads from the EPA flux analysis, with the storm event concentrations as an additional consistency check of the time-interval model predictions.

However, the model results sometimes differ from the data. For example, for the May 1999 storm data collected at Coltsville and West Branch (Figure 2-25), the simulated peak TSS concentrations are on the order of 100% greater than the measurements, as noted by the Reviewer. For this particular event, the simulated flow is lower than the measured flow; therefore, the resulting simulated loading (i.e., flow times concentration) for the event is closer to the measured loading than the predicted TSS concentrations indicate.

The significance of these differences for purposes of overall model calibration are impossible to assess since the upstream inputs of sediment used during the calibration runs relied on empirical data rather than the HSPF output. Since more than 75% of the stream associated input to the study area crosses this boundary this leaves HSPF inaccuracies affecting only 25% of the inputs an effective further diminished by the fact that these tend to be distributed along the length of the study area. Some consideration and subsequent discussion of this issue is recommended.

RESPONSE 3-FB-4:

EPA had considered this issue as part of the model calibration process, and agrees with the Reviewer that the fate and transport model is relatively insensitive to solids loadings predicted by HSPF. The predictions from HSPF provide a means of estimating solids loadings from the contiguous drainage area that would be difficult to derive from data. The effect of uncertainty in these estimates will be considered and subsequently discussed in the Model Validation Report.

In addition to consideration of high frequency TSS characteristics the presentation of HSPF would benefit from a more detailed discussion of the methodology leading to the specification of water temperature, a parameter passed to the bioaccumulation model. It appears that a moderately complex heat calculation is performed but details are not included in the calibration reports. A quick search indicates that they are also limited in the MFD.

RESPONSE 3-FB-5:

In HSPF, water temperature is modeled by performing an energy balance in each stream segment; the HTRCH submodule within the RCHRES module of HSPF performs these calculations. Heat and energy inputs to the stream are determined from the temperature of non-point, point, and boundary inflows, and from the following meteorological data: solar radiation, air temperature, dew point temperature, wind speed, and cloud cover.

Water temperature and heat content (in BTUs) of surface runoff and interflow (calculated in the PERLND and IMPLND modules) are estimated from air temperature using the following linear regression equation:

$$SLTMP = ASLT + BSLT*AIRTC$$

where:

SLTMP = Soil layer temperature (°C)

ASLT = Y-intercept (°C)

BSLT = Slope (-)

AIRTC = Air temperature (°C).

For the Housatonic River watershed model, the air temperature is input from various gages on an hourly time step, and the ASLT and BSLT parameters are varied monthly to represent seasonal patterns.

The lower soil layer temperature and the groundwater layer runoff temperature (from which the groundwater heat contribution is determined) are user-defined, and are usually derived from local shallow groundwater temperatures. These temperatures are also specified on a monthly basis to represent seasonal variability.

For each of the stream reaches in the watershed model, the HTRCH submodule performs the energy balance and estimates the stream water temperature. HTRCH accounts for inputs and outputs of heat in a reach through three heat-transfer processes: (1) heat transfer by advection into or out of the reach; (2) heat transfer across the air-water interface; and (3) heat transfer across the sediment-water interface. Heat is considered to be a thermal concentration that is uniform within a reach and assumed to advect at the same rate as the stream flow. The net heat exchange (Q_T - kcal/m^2/interval) at the water surface and sediment interfaces is represented by the following heat balance equation:

$$Q_T = Q_{SW} + Q_{ATM} + Q_B + Q_H + Q_E + Q_P + Q_G$$

where the flux terms are as illustrated in Figure 3-FB-5.

Radiational energy transfers at the water surface are estimated from solar radiation (short-wave) and cloud cover and temperature (longwave) data. Evaporative transfers are determined from wind, air temperature, and dew point temperature data. Conduction/convection transfers are determined from air temperature and wind. Energy transfers between the sediment and the water column are estimated from sediment temperature, which is user-defined, and is varied on a monthly basis to allow for seasonal variations.

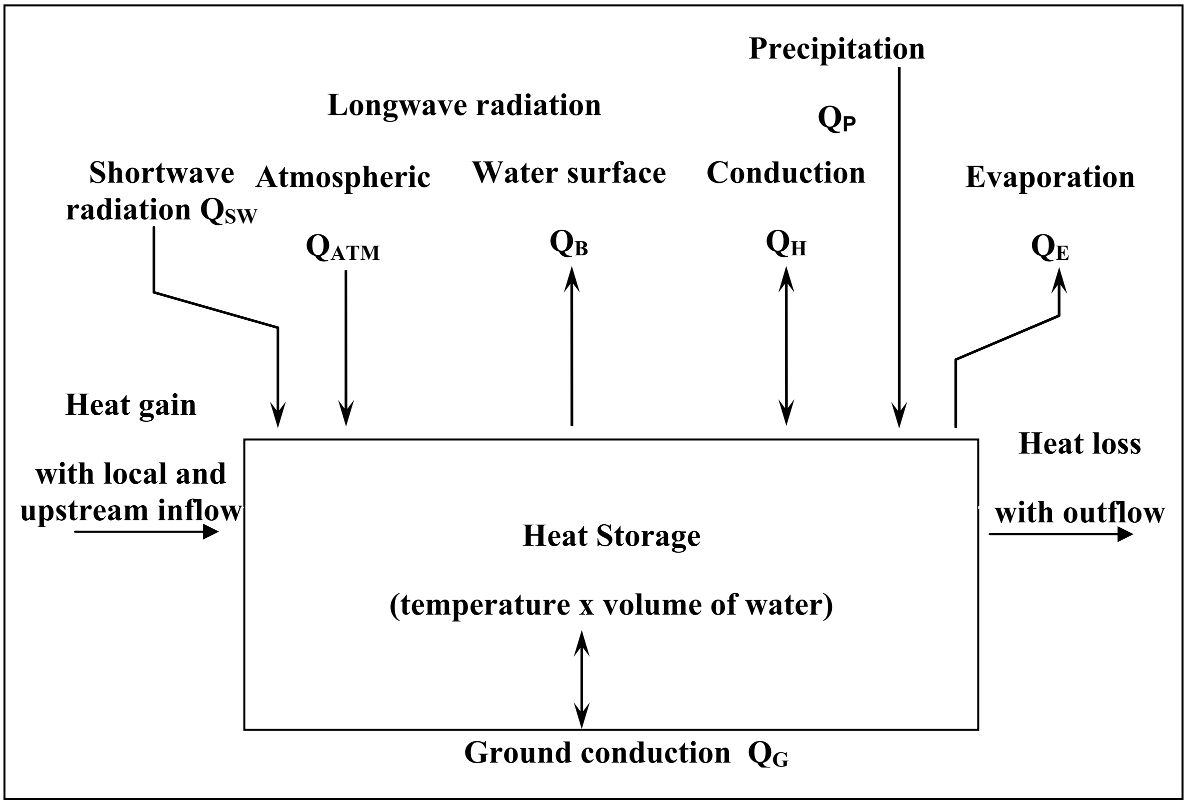


Figure 3-FB-5 Heat Exchange Processes Represented in the HTRCH Submodule of HSPF

The basic equations for each of the heat flux components were derived from a study of heat transfer processes between the atmosphere and water surfaces by TVA (1972). The formulations for each component are described below:

The shortwave radiation is estimated by the equation:

$$Q_{SW} = 0.97 * CFSAEX * SOLRAD * 10$$

where:

0.97	=	Fraction of incident radiation absorbed
CFSAEX	=	Ratio of radiation incident to water surface to measured radiation; includes shading
SOLRAD	=	Solar radiation (Langleys/interval)
10	=	Conversion factor from Langleys to kcal/m^2

The longwave radiation components are estimated by:

$$Q_{SW} + Q_B = SIGMA * [(KATRAD * 10^{-6} * CLDFAC * TAKELV^6) - TWKELV^4] * DELT60$$

where:

SIGMA = Stephan-Boltzman constant multiplied by 0.97 to account for emissivity of water

TWKELV = Water temperature (°K)

KATRAD = Atmospheric longwave radiation coefficient

CLDFAC = $1.0 + (0.0017 * CLOUD^2)$

CLOUD = Cloud cover, expressed as tenths (range = 0 - 10)

TAKELV = Air temperature corrected for elevation (°K)

DELT60 = Time interval (minutes) divided by 60

The conduction/convection heat transport is estimated by the following relationship, dependent on the difference between the water temperature and the air temperature:

$$QH = CFPRES * (KCOND * 10^{-4}) * WIND * (AIRTC – TW)$$

where:

CFPRES = Pressure correction factor (dependent on elevation)

KCOND = Conductive-convective heat transport coefficient

WIND = Wind speed (m/interval)

TW = Water temperature (°C)

AIRTC = Air temperature (°C)

The evaporative heat transfer is estimated by:

$$Q_E = HFACT * (KEVAP*10^{-9}) * WIND * (VPRESA - VPRESW)$$

where:

HFACT = Heat conversion factor (latent heat of vaporization multiplied by density of water)

EVAP	=	Quantity of water evaporated (m/interval)
KEVAP	=	Evaporation coefficient
WIND	=	Wind movement 2 m above the water surface (m/interval)
VPRESW	=	Saturation vapor pressure at the water surface (mbar)
VPRESA	=	Vapor pressure of air above water surface (mbar)

The heat content of precipitation is determined by assuming that the rainfall has the same temperature as the water on which it falls.

The heat movement between the sediment and the overlying water is based on heat transfer through a surficial layer (consisting of water-saturated sediment) overlying the deeper layer, which is at an equilibrium temperature). Heat fluxes between the deeper layer and surficial layer, and between the surficial layer and water are computed, as well as sediment and water temperatures. The heat transfer between the deeper layer and surficial layer is computed as follows:

$$QGRMUD = KGRND * (TGRND - TMUD)$$

where:

QGRMUD	=	Heat transfer from deeper layer to surficial layer (kcal/m^2/interval)
KGRND	=	Deeper layer – surficial layer heat conduction coefficient (kcal/m^2/°C/interval)
TGRND	=	Equilibrium deeper layer temperature (°C)
TMUD	=	Surficial layer temperature (°C)

This heat transfer is used to update the surficial layer temperature as follows:

$$TMUD = TMUD + QGRMUD/CPR/MUDDEP$$

where:

CPR	=	Heat capacity of surficial layer (CPR assumed to be equal to heat capacity of water)
MUDDEP	=	Thickness of surficial layer (m)

The new surficial layer temperature is used to compute the heat transfer between the sediment and the water column as follows:

$$Q_G = KMUD * (TMUD - TW)$$

where:

KMUD = Sediment-water column heat conduction coefficient (kcal/m^2/°C /interval)

TW = Water temperature (°C).

Reference:

TVA (Tennessee Valley Authority). 1972. *Heat and Mass Transfer Between a Water Surface and the Atmosphere*. Tennessee Valley Authority, Norris, TN. Laboratory Report No. 14, Water Resources Research Report No. 0-6803.

For the most part I'd believe that water and air temperatures are equal through much of the study area. Where deviations might occur however, is in the deeper, lower energy backwater and/or pond areas. Are these areas stratified? How is that to be handled?

RESPONSE 3-FB-6:

For food chain modeling, the HSPF linkage is used to represent water temperatures within the PSA. As the Reviewer notes, water temperatures are not widely spatially variable over main channel PSA reaches. The Reviewer is correct that temperature deviations can occur in shallow, low energy environments. For this reason, FCM estimates of temperature for Reach 5D backwaters were adjusted based on site-specific measurements in the main channel and backwaters (R2 Resource Consultants Inc., 2002). Daily temperature data were collected between May 2 and October 10, 2001 for two sampling locations in the mainstem of Reach 5C and one sampling location in the Reach 5D backwaters.

These data indicated that backwater temperatures are two to three degrees warmer than Reach 5C temperatures during the summer, with no consistent differences during cooler periods. The linear relationship shown in Figure 3-FB-6 was applied to temperatures greater than 10 °C; backwater temperatures were assumed to be equal to Reach 5D temperatures during the winter season.

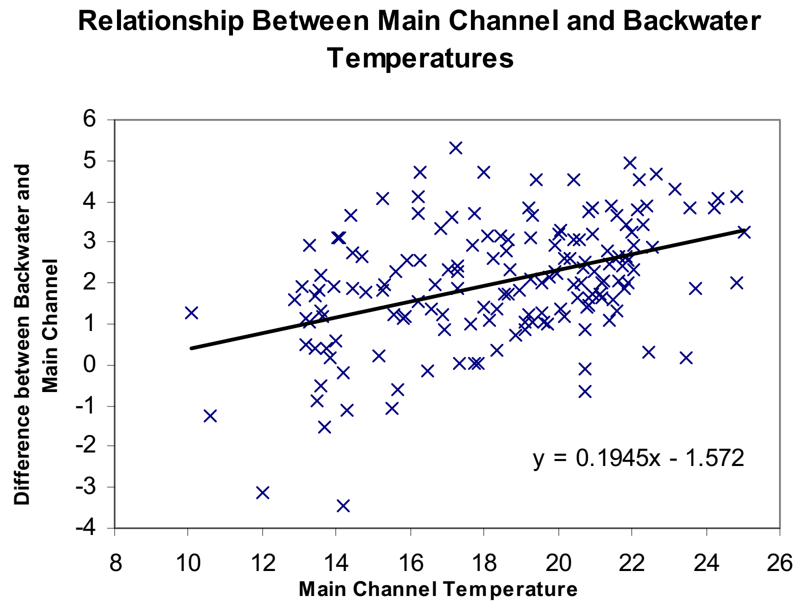


Figure 3-FB-6 Temperature Deviations Between Main Channel Housatonic River and Large Backwater Areas (Reach 5D)

Thermal stratification in the Reach 5D backwaters was not considered in the conceptual model. The depth of the backwater areas (typically, a few feet or less) is not sufficient to result in stratification over a vertical spatial scale relevant to fish bioaccumulation.

Reference:

R2 Resource Consultants, Inc. 2002. *Evaluation of Largemouth Bass Habitat, Population Structure, and Reproduction in the Upper Housatonic River, MA.* July 23, 2002.

The discussion of water temperature suggests that the water column throughout the study area is always well mixed. If this is the case this should be stated. If not the case, how will the water temperature near-bottom in Woods Pond be specified for bioaccumulation purposes?

RESPONSE 3-FB-7:

The Reviewer is correct that the water column was assumed to be well-mixed within each modeled reach. Therefore, the HSPF linkage was used to represent the temperature relevant to Woods Pond biota within FCM.

Woods Pond contains an area that is up to 16 feet deep, but the majority of the pond is 1- to 3-feet deep. The shallow banks of the pond provide extensive cover, including submerged macrophytes and woody debris, characteristics that make the habitat suitable for many resident fish species, particularly largemouth bass. As a result, the vast majority of fish habitat within Woods Pond consists of water depths over which vertical temperature stratification would not occur over a spatial scale relevant to fish bioaccumulation.

Although small vertical differences in water temperature likely occur in all study reaches, these differences are not sufficiently large to warrant quantitative representation in the model. The uncertainties in the temperature-dependent respiration equations within FCM are larger than the uncertainties in the temperature inputs to the model.

The discussion of HSPF also provides no indication of how the model calculates instream sediment transport. The majority of the streams entering the study area appear to be treated as simply conduits for the transport of sediment supplied by surface water drainage across sections of the watershed. Is this so? This has implications with regard to the amount of sediment delivered, its quality and the timing of delivery.

RESPONSE 3-FB-8:

In HSPF, sediment transport calculations are performed in each of the 65 HSPF stream reaches shown in Figure 2-6 of the Model Calibration Report, and listed in Table A.3-6 of Appendix A. Only 16 of these stream reaches are on the mainstem of the Housatonic River within the PSA. Thus, all the streams feeding the PSA as upstream boundary conditions and local tributaries are modeled with the SEDTRN module of HSPF, and therefore, are not simply conduits for the land-derived TSS loadings. Scour, deposition, and transport are represented in each of the stream reaches within the watershed model, although not at the level of detail as in EFDC.

Model Calibration Report Sections 2.5.1 and A.5.2.1.2 provide brief overviews of the solids simulation capabilities within HSPF and discuss the solids parameterization and calibration process. In response to the Reviewer's comment, there is additional discussion available that provides details on the instream solids transport formulations. This information is available in a recent technical paper by Donigian and Love (2003) and the HSPF User's Manual, available for download from EPA (www.epa.gov/waterscience/basins/bsnsdocs.html) and USGS (http://water.usgs.gov/software/code/surface_water/hspf/doc/hspfhelp.zip).

Reference:

Donigian, A.S., Jr. and J.T. Love. 2003. Sediment calibration procedures and guidelines for watershed modeling. Water Environment Federation TMDL 2003 Specialty Conference, November 16-19, 2003, Chicago, IL. WEF Specialty Conference Proceedings on CD-ROM (available at www.aquaterra.com/publications).

Moving next to EFDC beginning with the hydrodynamic model. Here I'm completely at a loss to explain why so much emphasis is placed on stage/discharge relationship and so little on the associated velocity field.

RESPONSE 3-FB-9:

Comparisons between simulated and measured stage and flow were developed using data collected within the 14-month calibration period. The emphasis placed on these data perceived by the Reviewer reflects the quantity of stage data collected during this time period. Compared to stage readings, fewer velocity measurements were made. The velocity data were used in model-data comparisons presented in the Model Calibration Report (see Section B.2.3.2.5).

The few direct current data presented in the report indicate that the model does a relatively poor job of specifying velocity despite its ability to define water level elevations. Review of the data presented indicated that this is most likely the result of two factors; the timing or phase of the stage along the study reach and/or the spatial segmentation used in the model.

RESPONSE 3-FB-10:

Comparisons between simulated and measured velocities (including velocities measured using the Acoustic Doppler Current Profiler (ADCP) were presented in Section B.2.3.2.5 of the Model Calibration Report. Because the velocity data were not collected within the 14-month calibration period, separate model simulations were conducted to develop these model-data comparisons. As described in Section B.2.3.2.5, the flow conditions at the time of the measurements were assigned in the model input, thus, the comparisons are not affected by timing or phase factors. The velocity data were obtained at Holmes Road, New Lenox Road, and Woods Pond Footbridge. The agreement between the model results and the data are directly related to the model representation of the constriction resulting from the bridge abutments. In the case of New Lenox Road, the distance between bridge abutments is reproduced well by the model grid and the agreement between the simulated and measured velocities is good. In the case of Woods Pond Footbridge, the constriction between the bridge abutments is, effectively, a sub-grid-scale effect that is not accounted for by the model grid. This is the cause of the underprediction of the velocities at that location.

The issues of phase shifts in model-data comparisons of stage and flow, and comparisons between simulated and measured velocities received careful attention as part of the hydrodynamic model calibration effort. The discussion of the hydrodynamic model performance, presented in the Model Calibration Report, Appendix B, Section B.2.3.2, addresses the issues of comparisons between simulated and measured flow and stage on storm-event time scales (Figures B.2-20 through B.2-29) and longer time scales (Figures B.2-14 through B.2-18). Comparisons between simulated and measured velocities (Figure B.2-37) and stage-discharge relationships (Figure B.2-19) are discussed. Comparisons of model simulations and observations during two extreme flow events are also discussed. The overall hydrodynamic calibration, summarized in the cross-plots of simulated vs. measured flow and stage (Figure B.2-38) supports the assessment that the hydrodynamic model performance is reasonable.

In discussing the model results no mention is made of the fact that while the model does provide a reasonable simulation of water level magnitudes, the timing of these elevations often differs substantially from that observed. This produces substantial differences in free-surface slope affecting both absolute velocity as well as the structure of the boundary layer. Since it is these characteristics that ultimately go into the calculation of boundary shear stress any error at this point may have profound effect on subsequent sediment transport calculations.

RESPONSE 3-FB-11:

EPA agrees that the agreement between the magnitude of simulated and measured stage and flow is reasonable. Cross-plots of simulated vs. measured flow and stage (Figure B.2-38) include differences that result from phase shifts; however, regressions of simulated vs. measured flow and stage produce coefficients of determination (r^2) of between 0.96 and 0.99, indicating that nearly all of the variance in the data is explained by the model. In cases where phase differences were indicated, these were attributed to uncertainty in the timing of inflows, not transport through the PSA. Calibration efforts have shown that attempting to substantially affect the timing of the simulated hydrographs for higher flow conditions did indeed result in simulated stage-discharge relationships that were inconsistent with observations. Application of the model to extreme flooding events that occurred outside the calibration period was performed to provide additional tests of the performance of the hydrodynamic model. It has been concluded from these calibration efforts that the phase shifts are related to uncertainty in the timing of flow inputs to the model domain and not the friction in the system that would affect the free surface slope, velocity, and the bed shear stresses.

Without care in the discussion of these factors this reviewer finds it hard to believe that this model can accurately simulate sediment/contaminant transport in the study area. This situation is best corrected by a careful analysis of the current meter data and comparisons with modeled velocity and a reasoned presentation of the results. If the existing data are not sufficient for this task an additional field effort should be initiated.

RESPONSE 3-FB-12:

EPA believes that adequate velocity data are available to support the modeling study. A careful analysis was performed of the ADCP data and cross-sectional velocity profiles; this analysis is presented in Section B.2.3.2.1.

In addition to questions regarding the temporal characteristics of the stage along the study area, the differences between modeled and observed velocities could be simply the result of the relatively coarse spatial segmentation selected for use in the model. Use of 20m square grids within the main stem channel results in nearly the entire channel width being covered by a single cell. This coarse segmentation will not accommodate the lateral variations in flow known to exist in a meandering channel such as that found in many portions of the study area. As a result it's not surprising that the cell average velocity can differ substantially from values observed at discrete points across the channel. Examination of channel bathymetry, plan view contours and the associated sediment distributions indicates that a minimum of three grid cells should be used across the main stem channel in order to adequately simulate hydrodynamics for use within subsequent transport estimates.

RESPONSE 3-FB-13:

Please refer to the response to General Issue 4.

If this results in an unacceptable increase in computation times consideration should be given to alternatives to batch runs including separating the individual models and developing discrete runs based on a well defined series of scenarios.

RESPONSE 3-FB-14:

Regardless of the number of grid cells used, EPA does not believe that either approach proposed by the Reviewer is appropriate for the Housatonic River modeling study. Although the modeling team considered separating the individual models, this approach was rejected for several reasons. The feedback at each time step between sediment-transport-induced bathymetric changes and the hydrodynamic calculations was judged to be a feature that should be included in the modeling analysis. To retain this feedback, the hydrodynamic and sediment transport components need to be run together. Experience has shown that it is much more efficient to calibrate sediment transport and PCB transport concurrently because the gradient in sorbed PCB concentrations between the water column and bed sediment imposes an important constraint on the sediment transport calibration. To evaluate sediment transport effects on PCB transport, these component models were integrated.

Use of steady-state flow conditions to construct a well-defined series of scenarios that span extended periods of time will reduce run time, but it could also introduce cumulative errors that are propagated and become significant over the course of a long-term simulation. There is a tradeoff between accuracy and the number of discrete steady-state scenarios that can be used, with greater improvements in simulation time but decreased accuracy as the number of discrete flows (i.e., flow bins) used to represent the hydrograph is decreased. In addition, the "feedback" inherent in running a continuous simulation is lost, similar to that discussed above.

The Reviewer's recommendations would also impact the ability to evaluate variability and uncertainty, which the Reviewers stressed during deliberations. Use of a series of steady state flow scenarios would produce a reduction in the variability of results computed by the model. Adopting the recommended approaches for reducing the computational requirements of EFDC would also have the effect of increasing uncertainty in the results.

The combination of coarse segmentation and questionable hydrodynamics makes it hard to believe that the model being tested is able to accurately simulate sediment/contaminant transport in the study area. Adding in the questions raised by Dr. Lick regarding the transport formulations being used, the interpretation of the empirical data and the vertical segmentation of the sediment column only adds to this concern.

RESPONSE 3-FB-15:

The segmentation used to represent the channel and floodplain of the PSA was developed based on a balance between computational feasibility and adequate representation of the physical characteristics of the PSA. The modeling teams representing EPA and GE both believe that the segmentation resulting in approximately one grid cell across the river channel is adequate to achieve the goals of the modeling study. In addition, EPA disagrees with the Reviewer's characterization of the hydrodynamics as "questionable." Based on the agreement between simulated and measured hydrodynamic data, EPA believes that the hydrodynamic calibration is adequate to achieve the goals of the modeling study.

With regard to the questions raised regarding transport formulations, there are a number of formulations accepted by modelers in common practice representing transport in riverine systems, with no consensus on which formulation is most applicable to a particular system. There are five individual processes that, collectively, comprise the transport formulations:

- Deposition of non-cohesive solids.
- Deposition of cohesive solids.
- Erosion of non-cohesive solids.
- Resuspension of cohesive solids.
- Bedload transport of non-cohesive solids.

Deposition of non-cohesive solids is described by the formulations in van Rijn (1984), which are well established and have been used in numerous sediment transport models.

Deposition of cohesive solids is a complicated process that is affected by variations in, for example, mineralogy, grain size distribution, and suspended solids concentration. This process can be represented at different levels of complexity, ranging from constant settling velocity to complex flocculation models. The settling velocity used in EFDC is a function of the weighted average of the settling velocities for washload and suspended load. In response to a recommendation provided during the Calibration Peer Review to use Dr. Lick's simple flocculation model (Section 4.4 of the Course Notes), EPA's modeling team implemented this flocculation model within EFDC. Although the model was able to represent the data from Dr. Lick's Couette flocculator, it was not able to reproduce the Housatonic River data; therefore, the flocculation model will not be used.

The equations used to describe the erosion of non-cohesive solids were developed by Garcia and Parker (1991) based on comparisons between simulations and experimental and field observations. Particle sorting and hiding mechanisms that are responsible for bed armoring for non-uniform sediment mixtures are included in the formulation.

Resuspension of cohesive solids is highly site-specific, and generic equations similar to those discussed above for erosion of non-cohesive solids are not available for cohesive solids. The parameterization of resuspension of cohesive solids used in EFDC was based on analysis of erosion data collected in the Sedflume experiments on Housatonic River sediment. This analysis was presented in Attachment B.5 of the Model Calibration Report. The Sedflume data were also reanalyzed in accordance with the approach and formulations suggested by Dr. Lick in his review comments; however, the variability in the Housatonic River data was not explained by those equations and analytical approach. This analysis is presented in Response 3-WL-9.

Non-cohesive bedload transport was described by the modified Engelund-Hansen formulation, which has been used in other sediment transport models. The Engelund-Hansen formulation was selected from the formulations available within EFDC because it had a number of desirable characteristics, including the ability to represent the effect of variation in sediment grain size within an individual grid cell, and the parameterization could be assigned from available site-specific information.

EPA believes that the model, as modified in response to the Peer Reviewers' comments, is able to adequately simulate sediment/contaminant transport in the study area and will provide a useful tool to evaluate the relative performance of potential remedial alternatives. Please refer to General Issues 4, 6, and 7 for additional discussion on these topics.

References:

Garcia, M., and G. Parker. 1991. Entrainment of bed sediment into suspension. *J. Hydraul. Eng*. 117:414-435.

van Rijn, L.C. 1984. Sediment transport, Part II. Suspended load transport. *J. Hydraul. Eng.* 110(11):1613-1638.

How is it then that the model outputs dealing with TSS are in reasonable agreement with observations? This brings us back to the matter of the formulation of model boundary conditions and in particular how the modelers specify the boundary conditions at the confluence. As discussed above, these are based primarily on empirical data and can be assumed to be reasonably accurate. Given the short calibration period of 14 months it's likely that this input represents the majority of the sediment moving through the area. The bed representing a small and generally negligible source. With this possibility, inaccurate specification of shear stress and the subsequent sediment transport would have little effect on model results be they TSS or estimates of deposition in low energy areas.

RESPONSE 3-FB-16:

EPA disagrees with the Reviewer regarding the significance of the bed as a source/sink for suspended sediment. As shown in the simulations during major storm events, e.g., for the May 1999 event (Event 1) (see Figure B.3-46 of the Model Calibration Report), a significant mass of sediment is resuspended on the rising limb of the runoff hydrograph and deposited on the falling limb. In this event, both the erosion and deposition fluxes in Reach 5A were approximately one-third that of the sediment load at the upstream boundary. Although the erosion and deposition fluxes were approximately equal in Reach 5A, the deposition flux in Reach 5C was slightly more than twice that of the erosion flux. Similar results were seen for Event 9 (Figure B.3-47). Thus, the reasonable agreement between measured and simulated TSS (especially for Reach 5C) would not have been achieved if the erosional and depositional processes were not simulated within an acceptable margin of error.

This lead to the uncomfortable postulate that what had been produced by all this work was in fact little more than a very complicated advection-diffusion model dominated by average cross-sectional flows and particle settling velocities.

RESPONSE 3-FB-17:

As indicated in Response 3-FB-16, EPA disagrees with the Reviewer's assessment of the model. Considering the resuspension, deposition, and bank slumping/erosion processes that occur in the PSA, the Reviewer's assessment is invalid because these sources/sinks of sediment are not negligible, as has been documented in the Model Calibration Report, during and immediately after runoff events.

To correct this situation several steps are required. First, the postulate might be tested by a test of the sensitivity of the model to upstream boundary conditions. This was not included in the present series of sensitivity tests.

RESPONSE 3-FB-18:

Evaluations of model sensitivity to upstream boundary conditions were performed and the results were reported in Section B.4.4 of the Model Calibration Report. EPA agrees that an evaluation of the response of the model to the reduction of PCB inputs at the upstream boundary to zero would be a useful demonstration of model performance and will consider the inclusion of such a simulation in the Model Validation Report.

Next, whatever the sensitivity run suggests, the structure of the velocity field must be better accommodated in EFDC. This will begin with a decrease in grid size followed by adjustments in channel friction factors to reduce the time differences between observed and modeled stage.

RESPONSE 3-FB-19:

EPA disagrees that the structure of the velocity field needs to be represented in greater detail to achieve the goals of the modeling study. As discussed in Response 3-FB-11, minor time differences between observed and modeled flows do not impact the ability of the model to simulate longer term changes in PCB concentrations. Please refer also to the response to General Issue 4.

With the hydrodynamics under control attention must turn to the sediment transport aspects of the model. I second all of Dr. Lick's comments with the additional proviso that even the sedflume results must be applied with care due to the differences in spatial scales affecting the flow regime in the river versus the flume.

RESPONSE 3-FB-20:

Responses to Dr. Lick's comments are provided below. With respect to the application of Sedflume results, the developers and those who have used the Sedflume (McNeil et al., 1996; Lick et al., 2005) indicate that differences in the spatial scales of the flow regime in the river versus the flume are less important than the bottom shear stress present in each. Erosion experiments involving similar size particles conducted in large-scale flumes and the Sedflume device produced similar erosion rates versus bottom shear stress (Roberts et al., 1998), suggesting that the scale effects are accounted for in the calculation of the bottom shear stress.

References:

McNeil, J., C. Taylor, and W. Lick. 1996. Measurements of erosion of undisturbed bottom sediments with depth. *J. Hydraulic Engng.* 122:316-324.

Lick, W., J. Gailani, C. Jones, E. Hayter, L. Burkhard, and J. McNeil, 2005. The Transport of Sediments and Contaminants in Surface Waters. Short Course Notes, University of California, Santa Barbara. January 2005.

Roberts, J., R. Jepsen, D. Gotthard, and W. Lick. 1998. Effects of particle size and bulk density on erosion of quartz particles. *J. Hydraulic Engng.* 124:1261-1267.

This is where some additional field data such as bathymetric change and/or radionuclide based estimates of sedimentation rates come in providing a "weight of evidence" corroboration of the empirical measurements.

RESPONSE 3-FB-21:

EPA has collected data to describe bathymetric changes, and both EPA and GE have collected sediment cores for measurement of radionuclide profiles that have been used to estimate sedimentation rates. The results of the Phase 2 Calibration will be compared to these data.

This is also a subject that would benefit from an increase in calibration period.

RESPONSE 3-FB-22:

Please refer to the response to General Issue 2.

The ultimate accuracy of the sediment transport formulations would also benefit from inclusion of side bank erosion in the model. It appears that it was left from the initial calibration runs because of time considerations.

RESPONSE 3-FB-23:

Please refer to the response to General Issue 9.

If so this is another reason for an increase in the calibration period. I do not advocate using portions of the verification runs for calibration purposes. Too often, this is self defeating.

RESPONSE 3-FB-24:

Please refer to the response to General Issue 2.

In addition to the inclusion of the sidebanks I'd recommend additional discussion of the dynamics to be applied to the floodplains. At present, they appear to be being treated simply as sinks. I believe that they also have the potential to serve as sources particularly during immediate post-storm periods when rainfall-runoff may displace sediments freshly placed on vegetation and/or the adjoining soil surfaces. Such processes may become increasingly important as alternative source areas are eliminated.

RESPONSE 3-FB-25:

The conceptual model of the floodplain as a sink for PCBs is based on the data analyses, including the fraction of PCBs located in the PSA floodplain soil presented in the RFI Report (BBL and QEA, 2003) and by the spatial pattern of PCB concentrations in the floodplain, which is consistent with the pattern of flooding from out-of-bank flow events reported in the final MFD. PCB mass estimates indicate that approximately 90% of the total mass of PCBs in Rest of River (Confluence to Long Island Sound) sediment and floodplain soil is located in the PSA and over 60% of the PCB mass in the PSA is located in floodplain soil. If substantial amounts of contaminated sediment that had been freshly deposited on vegetation and/or the adjoining soil surfaces were returned to the river, the bulk of PCB contamination would be in downstream sediment, rather than in the floodplain soil.

However, under the appropriate conditions, solids can be eroded from the floodplain during model simulations. The transport equations for non-cohesive solids in the floodplain are the same as in the river channel. The equations for the cohesive solids are also the same, with specification of a higher shear stress for erosion that accounts for the effect of vegetation based on the Universal Soil Loss Equation (EPA, 1993).

References:

BBL (Blasland, Bouck & Lee, Inc.) and QEA (Quantitative Environmental Analysis, LLC). 2003. *Housatonic River – Rest of River RCRA Facility Investigation Report.* Prepared for General Electric Company, Pittsfield, MA.

EPA (U.S. Environmental Protection Agency). 1993. *PRZM-2, A Model for Predicting Pesticide Fate in Crop Root and Unsaturated Soil Zones: Users Manual for Release 2.0*. EPA Office of Research and Development, EPA/600/R-93/046. March 1993.

With the model including the proper range of dynamics and the range of source sink areas a mass balance calculation must be conducted to insure that numerical artifacts neither produce nor consume mass. There was no demonstration in the present reports that the models were tested for mass continuity. This should be considered for both sediment and PCBs.

RESPONSE 3-FB-26:

Conservation of mass is a basic model requirement. Tests for mass continuity have been performed for both PCBs and solids. The mass balance was performed over the course of the 14-month model calibration period. Solids and PCB mass were conserved; the differences were approximately 0.006% and 0.003% in solids and PCB mass, respectively, over the course of the simulation. These small differences are well within what is typically considered to represent mass conservation, indicating that the numerical techniques within EFDC are operating properly.

An additional factor missing from the sediment/contaminant model is the matter of PCB volatility. A variety of studies have shown that this factor can result in significant PCB flux (see Thibodeaux et.al, 2002 - ACS Symposium Series 806:130-149). It is not clear from the discussion provided why it was neglected in the present models.

RESPONSE 3-FB-27:

Please refer to the response to General Issue 5.

Douglas Endicott:

Much effort has been devoted to collecting a robust data set for model calibration. However, in any project of this complexity there are always opportunities for additional data collection to address important gaps in the dataset. Supplemental data can be critical in terms of strengthening the model calibration. The adequacy of the models could be strengthened in a number of ways, each involving the collection of additional data and other information:

Partitioning data:

Since the modeling team don't understand what is going on at 15% of the coring locations (many but not all are very-low organic carbon sediments), it would be appropriate to do some adsorption experiments using sediments that deviate from the equilibrium partitioning behavior.

RESPONSE 3-DE-1:

As discussed in the MFD (Section 7.5), EPA conducted special studies to gain a better understanding of the behavior of PCBs in sediment with low organic carbon content. Samples of sediment were collected from Reach 5A. These samples were examined by light microscopy, and surface coatings were observed on many of the quartz grains. Subsequently, the samples were analyzed by Jesse Roberts and Dr. Rich Jepsen at Sandia National Laboratories using scanning electron microscopy (SEM), X-ray diffraction (XRD), and energy-dispersive system (EDS) via microprobe. The coatings were observed in almost every sample and were found to be enriched in carbon and chlorine. Sandia interpreted these results to indicate that PCBs are present in this coating. EPA does not believe that additional experiments are necessary to explore this phenomenon further to achieve the goals of the modeling study. Please refer also to the response to General Issue 5.

Low-flow sediment-water flux:

Methods of measuring bioturbation activity and/or diffusive flux should be investigated and employed. Deploying benthic chambers is one option.

RESPONSE 3-DE-2:

The overall schedule for the modeling study does not allow for planning and conducting additional field studies, and EPA believes it is in the best interests of the public to move the Rest of River study forward as expeditiously as possible, consistent with achieving the goals of the program. In addition, EPA does not believe that additional studies are necessary to calibrate and validate a model that will be useful in evaluating remedial alternatives. However, supplemental analyses of bioturbation activity in the Housatonic River are being conducted that integrate freshwater bioturbation literature with site-specific observations of benthic community assemblages. Please refer also to the responses to General Issues 1 and 6.

Sediment mixing:

Experimental approaches that could be applied in the field should be investigated. At a minimum, the density of benthic invertebrates (including vertical distribution) could be measured in sediment core samples as a basis for evaluating mixing depth.

RESPONSE 3-DE-3:

Please see Response 3-DE-2.

Potential PCB sources other than sediment bed and transport across confluence:

Neglecting the WWTP, tributaries, and groundwater as PCB sources in the PSA is not justified by the available data. The WWTP effluent has apparently not been monitored for PCBs, which I think may be an unfortunate mistake. Rationale for this offered by EPA (no gradient in water or sediment near outfall) is not compelling. We have observed PCB concentrations of 20-30 ng/L routinely in untreated sewage throughout New York and New Jersey. If, for example, the Pittsfield WWTP effluent were to contain PCBs at 20 ng/L, that would constitute a source of 0.3 kg/yr to the river. That doesn't sound like a lot in comparison to what's flowing across the confluence currently, but what about after remediation upstream is completed?

RESPONSE 3-DE-4:

EPA did not neglect the WWTP, tributaries, and groundwater as potential sources of PCBs to the PSA. In addition to the information provided to the Peer Reviewers as part of the Modeling Framework Design, Model Calibration Report, and related Rest of River documents, EPA has been working with GE on all aspects of PCB cleanup associated with the facility and does not believe there are other known sources of PCBs that must be accounted for in the modeling study.

The lack of any measurable increase in PCB concentrations in either water or sediment immediately downstream of the Pittsfield WWTP discharge is not the only basis for the conclusion that the WWTP is not a source of PCBs to the river. The City of Pittsfield has conducted analyses of sludge samples and final effluent from the facility. Analysis of these samples was non-detect for PCBs.

Similar arguments apply for tributaries and groundwater. EPA defends no groundwater monitoring by stating that too many measurements would be required to meet modeling accuracy requirements. Is this a rationale for neglecting the process? I am not sure EPA's QAPP process recommends "do nothing" in this situation.

RESPONSE 3-DE-5:

EPA did not neglect tributaries and groundwater influx as potential sources of PCBs to the river. In response to a similar question regarding tributaries raised by the Reviewer prior to the Peer Review Public Meeting, EPA indicated that it was not necessary to consider potential contaminant contributions from tributaries because: (1) in its extensive characterization of potential contamination of residential/commercial properties in the Pittsfield area, MDEP did not find PCB contamination in the vicinity of the tributaries; (2) no other known or suspected sources of PCBs occur in the drainage basins of the tributaries to the PSA; and (3) investigations of surface water, sediment, and floodplain soil conducted by MDEP, GE, and EPA during the EPA Rest of River investigation support the conclusion that the tributaries are not a source of PCBs to the river.

In response to a similar previous question from the Reviewer regarding groundwater influx, EPA indicated that groundwater influx is included in the model as part of a lumped parameter termed “vertical flux” that combines the processes of advection and diffusion. In the Conceptual Site Model (Section 4.5.2 of the final Modeling Framework Design), the processes of advection and diffusion are discussed in some detail, and the lumped processes are concluded to be “. . . an important factor [for the] spatial distribution of water column PCBs. It was noted in that response that advective flux is considered to be a calibration parameter and synoptic bed sediment and water column data were available to calibrate the flux...”

Therefore, sampling of groundwater flux was not necessary because it did not explicitly need to be parameterized for the model. In addition, sampling of groundwater flux would have required a program of vast scope to adequately characterize the substantial variability expected in the hyperheic zone across space and time throughout the PSA. A Quality Assurance Project Plan (QAPP) is designed to ensure that data quality objectives (DQOs) for a study are achieved. Because of the model construct, data sufficient to explicitly characterize groundwater flux were not identified as a DQO.

PCB volatilization:

Neglecting PCB volatilization in EFDC may not be justified. The accuracy of PCB flux estimates being used to justify the neglect of volatilization as a loss process in the PSA, depend upon good values of Henrys constant. The best experimental data I am aware of was published by Holly Bamford (Bamford, Poster and Baker, J. Chem Eng. Data, 2000, 45, 1069-1074). She measured Hlc’s over a range of temperatures for numerous congeners, and also generalized the results into predictions for all of the PCBs. She found that PCB Henrys constants depended more on the number of ortho-chlorines than on the homolog. If I substitute one of her Hlc values (a representative congener at 18 degrees C) for the value used in the RFI, and repeat the volatilization rate calculation, I get a rate and flux that is about twice as large. I suggest the modeling team evaluate Bamford’s data and consider revising the PCB volatilization flux calculation accordingly.

RESPONSE 3-DE-6:

Please refer to the response to General Issue 5.

Partitioning of PCBs to biotic organic carbon:

This was lost with the departure of AQUATOX. Instead, we get the assumption that suspended solids (with seasonally-invariant organic carbon content) has the same PCB sorbent capacity as phytoplankton, periphyton, macrophytes, etc. This is an assumption that should be stated in the PCB transport and fate model, and justified by data. I am not convinced that POC and phytoplankton/ periphyton carbon are interchangeable as sorbents for PCBs. I would like to see the evidence from the PSA that supports this assumption.

RESPONSE 3-DE-7:

Site-specific data support the assumption that periphyton and water column suspended solids absorb approximately the same proportion of PCBs on an organic carbon-normalized basis. These media are the most important pools of organic carbon in the pelagic portion of the bioaccumulation model because they serve as the base of the food web (i.e., water column invertebrates are assumed to have a diet consisting of POM). Concentrations of PCBs in macrophytes are not considered in the linkage between EFDC and FCM because macrophytes do not comprise an important component of the aquatic food web for invertebrates and fish. Section 3 of Attachment C.14 of the Model Calibration Report discusses PCBs in POM data for the PSA as well as available plant/algae data (summarized in Table 3-DE-7 below).

Table 3-DE-7

Analysis of PCBs in Suspended Organic Matter and Biological Organic Carbon Pools (mg/kg OC)

Sample/Analysis	Sample Size	Minimum	Maximum	Central Tendency
PCBs in POM (measured – high volume filter samples)	4*	150	256	211
PCBs in POM (estimated using 3-phase partitioning)	196	16	1,135	117
Periphyton (on macrophytes)	4	186	494	334
Periphyton (on hard substrate)	5	119	781	453
Macrophytes	5	29	77	49.1
Filamentous algae	4	0.4	3.8	1.4

* Duplicates are averaged and counted as a single sample; high flow samples omitted from analysis.

Directly measured PCBs in POM taken from high-volume surface water samples had a mean tPCB concentration of 211 mg/kg OC. Measured concentrations in periphyton (on macrophytes) exhibited a central tendency 58% higher. However, the range of the samples measured (186 to 494 mg/kg OC) is similar to the range measured in POM (150 to 256 mg/kg OC). Furthermore, periphyton includes live biological material, including bacteria and various microorganisms. As such, a small degree of biomagnification from detritus to periphyton is expected.

In EFDC, it was not possible to represent differential PCB partitioning in algal and non-algal suspended solids and other organic carbon pools because the data required for these refinements were not available. Refinement of the partitioning formulation to explicitly represent biotic particles would be expected to be of limited importance in Reaches 5A, 5B, 5C, and 6, where the long term solids and PCB mass fluxes are mainly controlled by high-flow conditions (during which the contribution of algal solids is of limited importance). Although sorption to algal solids may be relatively more important in the backwater areas (Reach 5D), the degree of sorption to algal carbon does not necessarily warrant the abandonment of a generic carbon-normalized sorption model. This is consistent with the decision to pool organic carbon sources within the water column and the sediment bed that was influenced by Reviewer comments received during the initial Peer Review of the Model Framework Document (MFD) in 2000. As currently implemented in EFDC, the different carbon pools are considered in the representation of PCB partitioning to the extent that the presence of these carbon pools is reflected in the suspended solids data and water column partitioning data sets that were analyzed and/or modeled.

The 3-phase partitioning approach that was selected for use in EFDC has previously been used in a number of well-known and widely applied fate and transport models (e.g., WASP5 and EXAMSII; Ambrose et al., 1993; Burns, 1997). It is also consistent with the partitioning formulation that was originally described in the MFD. EPA believes that, in the absence of more detailed data and a model that includes the simulation of algal productivity, use of a more refined representation of partitioning is not justified.

References:

Ambrose, R.B., T.A. Wool, and J.L. Martin. 1993. *The Water Quality Analysis Simulation Program, WASP5, Version 5.10, Part A: Model Documentation*. EPA, Office of Research and Development, Environmental Research Laboratory, Athens, GA.

Burns, L.A. 1997. *Exposure Analysis Modeling System (EXAMS II), User's Guide for Version 2.97*. EPA Research Laboratory, Athens, GA. 133 pp.

Monitoring of flow, solids and PCB concentrations at the boundary condition:

I am concerned that monitoring the flow, solids and PCB boundary condition above the confluence has not been emphasized enough. This results in unacceptable uncertainty in the upstream boundary condition. I think EPA and GE should consider adding more continuous instrumentation along with the pressure transducers, including ADCP and/or transmissometers. These would improve the flow measurements and allow continuous TSS monitoring, and could be used to make more robust estimates of the boundary conditions. Boundary conditions will become increasingly important as river remediation moves forward.

RESPONSE 3-DE-8:

EPA currently conducts continuous in-stream turbidity monitoring, with grab samples (4-point daily composite) collected for TSS and PCBs twice monthly upstream and downstream of the area being remediated. Flow is derived from the USGS Coltsville gage; continual changes in channel configuration due to the ongoing remediation make in-stream measurements in the remediation area questionable. In recognition of these changing conditions, EPA has established an alternative upstream boundary for the model at Newell Street (upstream of all in-stream remediation) with monitoring for flow, TSS, and PCBs. The new configuration of the reach of river between the new boundary and the confluence is represented in the model, which will be used after validation to evaluate future conditions and remedial alternatives.

Streambank erosion:

It appears that a major component of the interaction between the floodplain and the river occurs via erosion of the streambank. Active undercutting of sediments deposited on the riverbank is evident in the upper half of the PSA. This mechanism should be included in the sediment and contaminant transport model during calibration, not after.

RESPONSE 3-DE-9:

Please refer to the response to General Issue 9.

Marcelo H. Garcia:

Overall, the model does account for all the relevant processes. One exception is stream bank erosion which is currently not explicitly included and could be a major source of PCB.

RESPONSE 3-MG-1:

Please refer to the response to General Issue 9.

While the mathematical model does account for such processes, the computational grid used for the main river channel is too coarse and prevents the calibrated model from resolving the spatial scales needed to assess important processes within the channel itself and along its margins.

RESPONSE 3-MG-2

Please refer to the response to General Issue 4.

Frank Gobas:

With regards to the EFDC model, the model contains several processes controlling the fate of PCBs. The key processes that are included in the model are sediment-water diffusive exchange, solids settling & resuspension and flow. There are some key processes that are acknowledged in the model architecture (Figure 5-1) but which are not fully considered in the application of the model.

For example, degradation in the sediment is not considered in the model. The authors state that the rate of dechlorination is too small to be significant. They base this conclusion on the lack of a change in CL:BP ratios between originally discharged and current Aroclor 1260 in Reach 5A. However, this ratio does decrease in the lower portion of Reach 5A (Figure 5.15), hence suggesting dechlorination.

RESPONSE 3-FG-1:

Degradation of sediment has been observed in Woods Pond (Bedard and May, 1996); however, the magnitude of dechlorination was determined to be limited (Bedard et al., 1997; Wu et al., 1997), and the process was justifiably excluded from quantitative consideration in the modeling study (BBL and QEA, 2003).

Figure 5-15 depicts (at most) a very weak trend of declining Cl:BP ratios across the PSA. The decision to exclude dechlorination from the model was based on multiple lines of evidence, including more rigorous assessments than those depicted in Figure 5-15. For example, EPA conducted an extensive analysis of PCB congener composition (Attachment C.18 of the Model Calibration Report) to evaluate the degree of compositional change that occurs across sample type and river reach. Overall, the congener evaluation indicated that although some differences in profiles can be seen, most media exhibit congener profiles similar to Aroclor 1260 across all reaches. The degree of dechlorination of Housatonic River PCB mixtures was not large, especially in comparison to PCB sites with lower chlorination levels, such as the Hudson River (Chen et al., 1988). In the evaluation of congener profiles and specific individual congeners, major pattern differences across the PSA were not typically observed. The congener evaluation yielded findings consistent with site-specific assessments of sediment conducted by BBL and QEA (2003) as well as Bedard and May (1996), Bedard et al. (1996, 1997), and Bedard and Quensen (1995). Active dechlorination could not be detected in Woods Pond sediment that was incubated for more than a year (Bedard et al., 1995). Furthermore, because most of the PCBs in the PSA were discharged to the river several decades ago, any historic degradation that has occurred should largely be reflected in the current exposure levels that are assigned as initial conditions in the model.

References:

Bedard, D.L., L.A. Smullen, K.A. DeWeerd, D.K. Dietrich, G.M. Frame, R.J. May, J.M. Principe, T.O. Rouse, W.A. Fessler, and J.S. Nicholson. 1995. Chemical activation of microbially mediated PCB dechlorination: a field study. *Organohalogen Compounds* 24:23-28.

Bedard, D.L. and J.F. Quensen III. 1995. Microbial Reduction Dechlorination of Polychlorinated Biphenyls. Chapter 4. In L.Y. Young (Ed.), *Microbial Transformation and Degradation of Toxic Organic Chemicals*. Wiley-Liss. pp. 127-216.

Bedard, D.J. and R.J. May. 1996. Characterization of the polychlorinated biphenyls in the sediments of Woods Pond: Evidence for microbial dechlorination of Aroclor 1260 in situ. *Environ. Sci. Technol.* 30(1):237-245.

Bedard, D.L., S.C. Bunnell, and L.A. Smullen. 1996. Stimulation of microbial para-dechlorination of polychlorinated biphenyls that have persisted in Housatonic River sediment for decades. *Environ. Sci. Technol.* 30:687-694.

Bedard, D.L., H.M. Van Dort, R.J. May, and L.A. Smullen. 1997. Enrichment of microorganisms that sequentially meta, para-dechlorinate the residue of Aroclor 1260 in Housatonic River sediment. *Environ. Sci. Technol.* 31:3308-3313.

Chen, M., C.S. Hong, B. Bush, and G.-Y. Rhee. 1988. Anaerobic biodegradation of polychlorinated biphenyls by bacteria from Hudson River sediments. *Ecotoxicology and Environmental Safety* 16:95-105.

Wu, Q., D.L. Bedard, and J. Wiegel. 1997. Temperature determines the pattern of anaerobic microbial dechlorination of Aroclor 1260 primed by 2,3,4,6-tetrachlorobiphenyl in Woods Pond sediment. *Applied and Environmental Microbiology* 63(12):4818-4825.

To shed more light on this issue it is beneficial to explore changes in PCB composition over time on a congener specific basis. Although dechlorination may be a slow process, it can have a significant effect on the overall fate of PCBs in the Housatonic River if other loss rates of PCBs from the River are also slow. The latter appears to be case since net loss of PCBs from the 5A to 6 reaches is very small when expressed as a fraction of the mass of PCBs present in the River. This means that the river's response time to changes in PCB loadings is very long, i.e. it takes a long time for sediment concentrations to respond to a new loading regime. In slowly responding systems, slow processes can have an impact on the overall response time of the PCB concentration in the system and even be rate limiting. In that light I recommend that the authors include the degradation rates of PCBs in the model, preferably on a congener specific basis.

RESPONSE 3-FG-2:

Please see Response 3-FG-1, above. The importance of simulating individual PCB congeners is greatest when the congener composition of the mixture changes significantly as a result of either PCB degradation (i.e., dechlorination, metabolism, or breakdown) or PCB transformation (i.e., selective enrichment or depletion of congeners governed by environmental partitioning). To this end, EPA conducted an extensive analysis of PCB congener composition (Attachment C.18 of the Model Calibration Report) to evaluate the degree of compositional change that occurs across sample type and river reach.

The lack of major pattern differences across the PSA indicates that congener-specific modeling is less important than for other PCB-contaminated sites.

I also recommend that the volatilization of PCBs to the atmosphere is considered in more detail as it may be a significant loss rate for PCBs in the River. While volatilization from the river may be small due to small surface area, this surface area is significantly increased during flooding events when particulate material and water are distributed over large areas of floodplain. When flooding subsides, these particulate materials will be in contact with air for considerable times and PCBs may volatilize.

RESPONSE 3-FG-3:

Please refer to the response to General Issue 5.

The decay of organic matter in suspended sediments is recognized in Figure 5, but it is not clear whether it is actually considered in the application of the model. In the Housatonic River, where PCBs have been associated in sediments for a long time, it is possible that as a result of the relatively rapid decay of organic matter compared to a slow desorption rate of PCBs to the water, organic carbon normalized PCB concentrations increase relative to the water concentration, causing a suspended sediment-water disequilibrium that affects the PCB concentration in the water available for respiratory uptake in biota. Evidence of this process has been observed in suspended and bottom sediments in some other system (Environ. Sci. Technol. 37(4): 735-741).

RESPONSE 3-FG-4:

EPA assumes that the Reviewer is referring to Figure 5-1 of the Model Calibration Report, which shows “decay” as a process relevant to dissolved contaminant and particulate contaminant. This figure was not intended to represent the decomposition of particulate organic matter, but instead the decay of contaminant that is sorbed to particulate organic matter. Modeling the decay of contaminant is a capability of EFDC, but one that was not applied as described in Section 5.2.3.3.5.

Regarding carbon decay, the EFDC model neither calculates the decay of organic carbon in suspended sediment nor considers the effects of such decay on PCB concentrations in water. In the water column, the fraction of organic carbon (f_{oc}) was established based on site measurements. This spatial variation in f_{oc} was maintained constant in time throughout the calibration. The partition coefficients derived from site data reflect the in situ condition of Housatonic River sediment. Thus, any degradation of organic matter that is taking place in the Housatonic River is inherently reflected in the bulk sediment and pore water data. EFDC does not model the organic carbon cycle, which would be necessary to explicitly represent the process referred to by the Reviewer. Instead, EFDC uses three-phase partitioning of PCBs to calculate the concentrations of PCBs in water, DOM, and POM. Site-specific data support this simple model of PCB partitioning, as described in the Model Calibration Report Section 5.2.3.3.1. Inclusion of a more complex model that would include suspended sediment-water disequilibria was considered. However, such a model was rejected on the basis that the simpler three-phase partitioning model matched site-specific data reasonably well. This is consistent with the decision to pool organic carbon sources within the water column and the sediment bed that was influenced by Reviewer comments received during the initial Peer Review of the MFD in 2000. At that time, several of the Reviewers expressed concerns regarding the overall complexity of the AQUATOX model (e.g., Dr. Lick, pages 3, 11; Mr. Endicott, page 3; Dr. Shanahan, pages 4-5, 7). Essentially, the Reviewers felt that the model was over-specified, with too many parameters and excessive model uncertainty. A specific review comment (Mr. Endicott) questioned the need for kinetic models of PCB partitioning to detrital and planktonic organic carbon on the basis of “seemingly excessive complexity.”

Gobas and Maclean (2003) describe some theoretical and empirical grounds for enrichment of contaminant concentrations in suspended sediment. EPA agrees that disequilibrium can occur between suspended sediment and the water column under certain conditions. However, there are several reasons why the simpler equilibrium model was retained for the PSA:

- The degree of disequilibrium between suspended sediment and water decreases as the octanol-water partition coefficient increases. Gobas and Maclean (2003; Figure 2) showed that deviations from chemical equilibrium occur in several systems (Lake Ontario, Lake Superior, Lake St. Clair, Lake Erie, Green Bay), particularly for log K_{OW} values below 6.0. However, at log K_{OW} of approximately 7.0 (i.e., representative of Housatonic River PCBs dominated by Aroclor 1260), the deviations from chemical equilibria were much smaller.

- The relationships between chemical disequilibrium and K_{OW} (i.e., fugacity ratio) were not identical among North American freshwater systems. Fugacity ratios of high-K_{OW} substances (log K_{OW} values above 7.0) were sometimes above 1.0 and sometimes below 1.0 (Gobas and Maclean, 2003). Therefore, there is uncertainty associated with a suspended sediment-water disequilibrium model that must be traded off against the increased level of mechanistic representation offered by such a model.

- The magnitude of chemical disequilibrium exhibits an apparent positive relationship with lake depth. The data "suggest that the deeper lakes i.e., Lake Ontario (86 m) and Superior (147 m) exhibit larger sediment-water disequilibria than the shallower Lakes Erie (19 m), St. Clair (9 m), and Green Bay (15 m)" (Gobas and Maclean, 2003). Much of the PSA has water depths typically less than 1 m; therefore, the degree of organic decomposition is probably low relative to deeper lakes with large surface-to-bottom decomposition times. Extrapolation from lake environments to the Housatonic River would have significant uncertainty.

- Distribution coefficients for hydrophobic contaminants between suspended sediment and water can be increased as a result of organic carbon decomposition. However, this process can be compensated by internal primary production of organic matter. In the shallow reaches of the Housatonic River, primary production can be significant during warm months. Accordingly, the organic carbon cycle and budget are complex and would require significant increases in model complexity (and associated model uncertainty) to simulate mechanistically.

- The effect of organic carbon decay is further complicated by the differences in partitioning to different types of organic carbon (i.e., refractory or labile nature of the carbon sources).

Overall, EPA believes that the uncertainties associated with the incorporation of organic carbon decomposition (and other aspects of the organic carbon cycle) are sufficiently large that they outweigh the benefits of increased model complexity. This conclusion echoes the general sentiments of the Peer Review Panel following deliberations on the Model Framework Document in 2001. At that time, substantial concerns were raised over the proposed modeling of multiple biological carbon types in AQUATOX.

Reference:

Gobas, F.A.P.C. and L.G. Maclean. 2003. Sediment-water distribution of organic contaminants in aquatic ecosystems: The role of organic carbon mineralization. *Environ. Sci. Technol.* 37(4):735-741.

In summary, the model is focused on the description of sediment dynamics and sediment:water partitioning of PCBs but does not fully explore several other fate controlling processes that, considering the slow temporal response of the system, may have a significant effect on the outcome of the model.

RESPONSE 3-FG-5:

During the specification of the conceptual model, fate-controlling processes were considered in detail on a process-by-process basis (please refer to Section 4.4 of the Modeling Framework Design). Processes such as sediment decay and biodegradation of congeners were considered and included if there was sufficient evidence that they would have a significant effect on model outcome. Lacking such evidence, processes were not included.

EPA agrees that "slow" processes can become rate-limiting in systems with slow overall response times. However, the importance of these processes must be assessed in the context of the model objectives (please refer to the response to General Issue 1). For the purpose of comparing remedial options, including natural recovery, fate-controlling processes are included only if they improve the reliability of model simulations (and the differences among outcomes for remedial alternatives).

Another significant limitation of the EFDC model is its inability to model PCB congeners. The representation of PCBs with average properties (e.g. Kow) can produce a significant error in the calculations of PCB concentrations. The modeling of PCBs in terms of total PCBs has merits but it is not a state of the art modeling methodology. The tPCB modeling becomes a limitation when the model results of the EFDC model are transferred to the FCM model and used to assess ecotoxicological effects. With regards to assessing the ecotoxicological effects of PCBs, the current practice relies on assessing risks of effects based on congener concentrations. I strongly recommend that the EFDC model conducts congener specific calculations that can take advantage of available congener specific physical-chemical and biological data.

RESPONSE 3-FG-6:

The human health and ecological risk assessments considered congener-specific toxicity data. There were few relevant studies identified in the scientific literature that evaluated the toxicity of congeners relative to studies that evaluated tPCBs and/or Aroclors. Therefore, the emphasis in the risk characterizations was, in general, on tPCBs, and the interim media protection goals (IMPGs) will also be based primarily on tPCBs, with some consideration of toxic equivalence (TEQ).

Given this focus, the modeling study appropriately emphasizes simulation of the transport and fate of tPCBs using site-specific values empirically derived for such parameters as K_{ow} and BSAF. Such values implicitly integrate the properties of the congener mixture present at the site.

The lack of major differences in congener patterns across the PSA indicates that congener-specific modeling is less important than for other PCB-contaminated sites. The ecotoxicological effects of the PCB mixtures can still be assessed within the model framework by distributing the tPCB mass to individual congeners at the end of EFDC simulations, as is currently performed in the EFDC-FCM linkage. Please refer also to the response to General Issue 5 and Responses 3-FG-1 and 3-FG-2 above.

The bioaccumulation model contains the key processes controlling the uptake and elimination of PCBs in fish and invertebrates. Uptake from water and diet are included along with elimination to water and other excretion processes and growth dilution. The model also includes a de facto mechanism for biomagnification, apparently through a resistance factor CR that applies to the gill elimination rate but not to the gill uptake rate. The fact that the resistance does not apply to both uptake and elimination for this reversible process is not correct in my view. However, this practice produces a biomagnification effect that is not explicitly included in the model. The resulting model with the resistance factor can be expected to work well as it appears to do.

RESPONSE 3-FG-7:

This issue refers specifically to the use of the parameter c_R in the bioaccumulation model. This parameter scales the calculation of the PCB elimination rate. It was developed in response to the apparently contradictory observations that (1) laboratory experiments indicate that the exchange of PCBs across the gill surface is generally rapid, although (2) the rates at which adult field-collected fish eliminate PCBs are relatively slow. The parameter c_R was added to the model to account for the slow elimination rates. It is a simple multiplier of the computed gill elimination rate and fulfills this function.

The process of exchange across the gill is diffusive; therefore, a reduction in the elimination rate should properly be matched by an equal reduction in uptake. There is literature available that suggests that the limiting process in PCB uptake and elimination across the gill is not transfer across the gill, but transfer between blood and fat. The basis for this is described in the documentation of the Hudson River bioaccumulation model developed by QEA (QEA, 1999). Furthermore, a pharmacokinetic model of PCB uptake in fish has been developed (Nichols et al., 1990, 2004a, 2004b) that explicitly models blood flow to body tissues and exchange between blood and body tissues. It supports the conclusion that limited perfusion of fat tissues is the primary mechanism causing reduced elimination of PCBs in fish.

Simulation of the kinetics of transfer between blood and fat would require a pharmacokinetic model, i.e., one that explicitly models individual body tissues and the kinetics of exchange between them. FCM does not model the kinetics of transfer of PCBs among body tissues. Therefore, the parameter c_R was introduced as an empirical means of accounting for the observed slow elimination rates. This parameter affects only elimination, not uptake. Nonetheless, this has been sufficient for past applications of the model because uptake across the gill was typically not a significant source of contaminant. Uptake from food was much more important.

To confirm the assumption that contaminant uptake across the gill is a less important vector for bioaccumulation relative to uptake from food, EPA re-ran the linked model with the resistance factor c_R applied to both the gill uptake rate and the gill elimination rate. In the modified model, gill uptake was calculated as follows for fish species:

$$GillUptake \quad K_{ui}c(C_R)$$

where:

GillUptake	=	Gill uptake of contaminant µg/g-wet day.
K_{ui}	=	Gill uptake rate constant for species i (L/g-wet day).
c	=	Concentration of PCBs dissolved in the water (µg/L), where “dissolved” is defined to mean truly dissolved or bioavailable concentration.
c_R	=	Resistance factor accounting for PCBs in deep storage compartments (unitless).

The results are shown in Figure 3-FG-7, which is comparable to Figure 6-6 of the Model Calibration Report. The differences between Figure 3-FG-7 and Figure 6-6 are minor. Therefore, the approximation of the biphasic elimination process in the FCM kinetics is not appreciably affected by the inclusion or exclusion of the resistance factor in the model’s uptake kinetics.

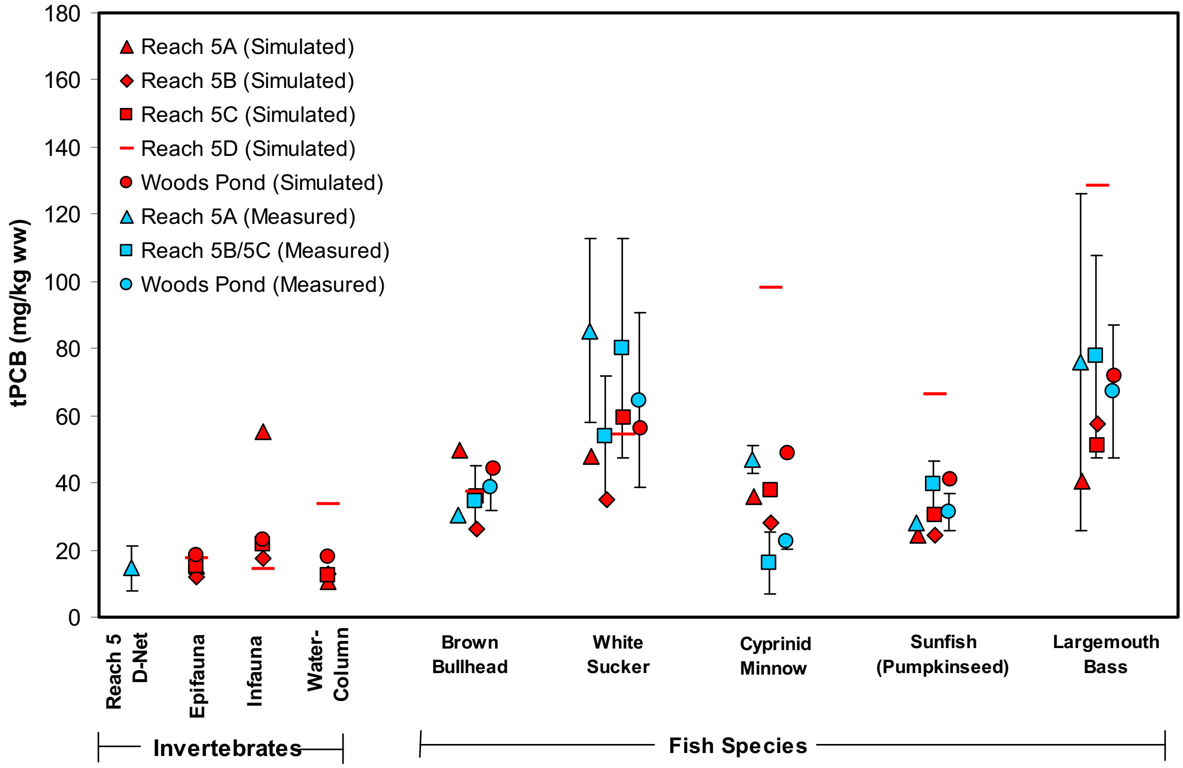
References:

Nichols, J.W., J.M. McKim, M.E. Andersen, M.L. Gargas, H.J. Clewell III, and R.J. Erickson. 1990. A physiologically based toxicokinetic model for the uptake and disposition of waterborne organic chemicals in fish. *Toxicology and Applied Pharmacology* 106:433-447.

Nichols, J.W., P.N. Fitzsimmons, F.W. Whiteman, T.D. Dawson, L. Babeu, and J. Juenemann. 2004a. A physiologically based toxicokinetic model for dietary uptake of hydrophobic organic compounds by fish. I. Feeding studies with 2,2',5,5'-tetrachlorobiphenyl. *J. Toxicological Sciences* 77:206-218.

Nichols, J.W., P.N. Fitzsimmons, and F.W. Whiteman. 2004b. A physiologically based toxicokinetic model for dietary uptake of hydrophobic organic compounds by fish. II. Simulation of chronic exposure scenarios. *J. Toxicological Sciences* 77:209-229.

QEA (Quantitative Environmental Analysis, LLC). 1999. *PCBs in the Upper Hudson River*. Volume 2 - *A Model of PCB Fate, Transport, and Bioaccumulation*. Report prepared for the General Electric Company, Albany, NY. Amended July 1999. URL: http://www.ge.com/files/usa/en/commitment/ehs/hudson/Volume_2.pdf



Notes: Simulated values represent FCM model simulations corresponding to the average age of all fish observed in the field (ages rounded to nearest whole number).

Measured values represent mean concentrations of tPCB (± 2 standard errors), for all fish within each combination of organism type and river reach.

Figure 3-FG-7 Simulated Versus Measured PCB Concentrations in Housatonic River Food Web Using Linked Model and Applying Resistance Factor to Both Gill Uptake and Elimination Rates

There are some processes that could be included such as egg and sperm deposition for spawning fish. However, I do not recommend this. The model is calibrated to quite a significant extent and adding further parameters that are included in the calibration recipe makes the model less transparent while any improvements in predictability are unlikely to be significant.

RESPONSE 3-FG-8:

EPA agrees that egg and sperm deposition for spawning fish should not be included in the model. Although maternal transfer of PCBs to eggs has been documented, there are inadequate site-specific data to assess the performance of a model that includes mechanistic representation of these processes.

As described in Response 2-EA-1, EPA does not agree that FCM has been calibrated "to quite a significant extent." However, EPA agrees that addition of further parameters/processes that are difficult to validate would increase the complexity of the model without improving the predictive value of the model.

A gap, perhaps in the reporting only, concerns the model for accumulation in aquatic macrophytes and algae. A simple lipid-water type partitioning model is unlikely to be successful in describing the bioaccumulation of PCBs in algae & macrophytes. Adding this component to the model may not have a big effect on the model outcome given the apparently strong linkage of the food-web to the sediment. However, it is important to ensure that the reporting of the modeling approach is complete.

RESPONSE 3-FG-9:

As applied to the Housatonic River, FCM does not explicitly represent uptake in algae or macrophytes. The base of the food web is defined to consist of particulate organic matter in the water column and organic matter in the sediment bed. Although organic matter in each of these compartments is composed of a variety of organic carbon types, the model does not explicitly differentiate among carbon sources (i.e., living biological material versus detritus, or refractory versus labile carbon sources).

The decision to pool organic carbon sources within the water column and the sediment bed was influenced by Reviewer comments received during the initial Peer Review of the Model Framework Document (MFD) in 2000. At that time, several of the Reviewers expressed concerns regarding the overall complexity of the AQUATOX model (e.g., Dr. Lick, pages 3, 11; Mr. Endicott, page 3; Dr. Shanahan, pages 4-5, 7). Essentially, the Reviewers felt that the model was over-specified, with too many parameters and excessive model uncertainty. A specific review comment (Mr. Endicott) questioned the need for kinetic models of PCB partitioning to detrital and planktonic organic carbon on the basis of "seemingly excessive complexity." The level of complexity that was included in AQUATOX was subsequently reduced through the use of the FCM bioaccumulation model.

The resulting assumption implicit in FCM is that food sources for invertebrates and fish contain (on average) approximately the same quantity of contaminant (on an organic carbon-normalized basis) as do suspended sediment and bottom sediment. Response 3-DE-7 indicates that periphyton and particulate organic matter have similar PCB concentrations. Other media, such as filamentous algae and macrophytes, appear to have different bioaccumulation potential but are not important components of the conceptual model food web.

While not a process, I question the wisdom of not including some other target species in the model such as muskrat, waterfowl and raptors. These organisms are susceptible to high concentrations of PCBs due to bioaccumulation and "dose-response" relationships exist for risk analysis purposes. This may have been addressed in an earlier bounding exercise.

RESPONSE 3-FG-10:

The Ecological Risk Assessment considered eight Assessment Endpoints. These Assessment Endpoints were selected to represent the spectrum of susceptible and relevant ecological receptors in the PSA. Waterfowl and raptors were represented within these Assessment Endpoints; extensive work modeling exposures and effects for these taxa is presented in the Ecological Risk Assessment. IMPGs for these ecological receptors can be calculated for the aquatic components of the diet from contaminant concentrations predicted in the tissues of taxa that are explicitly modeled in FCM. With regard to muskrat, PCBs do not transfer appreciably from sediment to the types of aquatic vegetation preferred by muskrat, and there is no reason to expect that muskrats, or any other strict herbivore, would be at greater risk than the Assessment Endpoints (such as piscivorous mammals) that were selected; therefore, herbivorous mammals were not evaluated. Muskrat were commonly observed in the PSA during the ecological characterization.

Reference:

WESTON (Weston Solutions, Inc.). 2004. *Ecological Risk Assessment of the General Electric (GE)/Housatonic River Site, Rest of River*. Prepared for U.S. Army Corps of Engineers and U.S. Environmental Protection Agency.

Wilbert Lick:

In water quality models, the values for many parameters are commonly determined by parameterization, i.e., by varying the values of each parameter until the solution, however defined, fits some observed quantity. There are serious difficulties with this type of procedure. As a simple example, consider the erosion and deposition of sediments. In this case, a limiting situation is where there is a local steady-state equilibrium between erosion and deposition. Denote the erosion rate by E and the deposition rate by pw_sC, where p is the probability of deposition, w_s is the settling speed of the particles, and C is the suspended solids concentration. Local steady-state equilibrium then implies that

$$E = pw_sC \qquad (1)$$

Rearranging, one obtains

$$C \quad \frac{E}{pw_s} \qquad (2)$$

From this, it is easy to see that a numerical model can "predict" the observed value for C with an almost arbitrary value of E, as long as pw_s is changed accordingly, i.e., such that $E/pw_s = C$. For example, a particular value of C can be obtained by high values of erosion and deposition or by low values of erosion and deposition, as long as they balance to give the observed value of C. For a predictive model, the values of E and pw_s can not both be determined from calibration of the model by use of the suspended solids concentration alone.

RESPONSE 3-WL-1:

Please refer to the response to General Issue 7.

As a practical illustration of this problem, consider the sediment and contaminant transport modeling in the Fox River (Tracy and Keane, 2000). Two groups independently developed transport models. Each group calibrated their model based on suspended sediment concentration measurements. Each group believed that the parameters used in their model were reasonable.

However, the results predicted by the two models were quite different, both in the transport of sediments and of contaminants. As an example, the amount of sediment resuspended at a shear stress of 1.5 N/m^2 (a large but not the maximum shear stress in the Fox) was predicted by one group to be 11.3 g/cm^2 (on the order of 10 cm) while the other group predicted 0.1 g/cm^2 (or 0.1 cm), a difference of two orders of magnitude.

This difference of course has a direct impact on the choice of remedial action. Small or no erosion at high shear stresses indicates that contaminants are probably being buried over the long term and natural recovery is therefore the best choice of action. Large amounts of erosion indicate that buried contaminants may be uncovered, be resuspended, and hence will contaminate surface waters; dredging or capping is therefore necessary. The differences in the model estimates by the two groups make it difficult to decide on the appropriate remedial action.

This may seem like a long and tedious introduction, but I want to make the point that models with many unconstrained parameters and especially models which include processes that are not described correctly as far as their functional behavior is concerned can lead to non-unique solutions; these can then lead to the incorrect predictions of long-term behavior.

RESPONSE 3-WL-2:

Please refer to the response to General Issue 7.

For the long-term prediction of sediment and contaminant fluxes, it is essential that the functional behavior of the most significant processes be described correctly. In this regard, the most significant processes are sediment erosion/deposition (including bank erosion and slumping) and the diffusional flux of contaminants between the bottom sediments and the overlying water. Erosion/deposition is significant not only because of the contaminants transported with the sediments, but a major question is whether erosion/deposition during big events will expose deeply buried contaminated sediments **or** deeply bury surficial contaminated sediments. Both are possible during big events and will strongly influence the contaminant flux in the future. Accurately predicting suspended sediment concentrations under present conditions, although necessary, is not sufficient for an accurate, long-term prediction of contaminant exposure and transport.

RESPONSE 3-WL-3:

Please refer to the response to General Issue 7.

As far as the diffusive flux between the sediments and overlying water is concerned, the magnitude of this flux is obviously important. However, for the long-term prediction, the depth over which this flux acts is also significant.

RESPONSE 3-WL-4:

Please refer to the response to General Issue 8.

I will concentrate my comments on the two processes of sediment resuspension/deposition and sediment-water flux of PCBs.

a. Sediment Erosion

The Shaker was developed and used in 1990-94 as a field device to measure the erosion potential of relatively undisturbed sediments in cores. No other device was available at that time. It was calibrated against the annular flume. Both the annular flume and Shaker measure net resuspension, i.e., resuspension of sediments in the presence of deposition. In contrast, Sedflume measures pure erosion, i.e., erosion of sediments with no deposition. Pure erosion is the quantity that is used in sediment flux equations and in water quality models.

In 1999 (Lick, et al., 1999; also see chapter 3 of class Notes), a comparison of the Shaker and Sedflume was made and the processes in each (as well as in the annular flume) were carefully examined. It was determined that, because of experimental artifacts in the annular flume (and hence inaccurate calibration of the Shaker), the annular flume and Shaker gave qualitatively correct results but did not give accurate quantitative results. Because of this, the use of the Shaker is not recommended.

However, the use of Sedflume is recommended, primarily to determine erosion rates as a function of shear stress and as a function of depth in the sediment. As a by-product, a critical shear stress for erosion as a function of depth is also determined. Sedflume should be used in conjunction with the Density Profiler (Gotthard, 1997; Roberts et al., 1998), which determines the bulk density of the sediments (including solids, water, and air) as a function of depth in the sediment core in a non-destructive manner. This allows the determination of sediment layering before erosion rates are measured (which is destructive) and allows the determination of erosion rates as a function of the bulk properties of the sediment in that layer. This has been done successfully in other places (e.g., the Kalamazoo River, McNeil et al., 2004 and Chapter 3 of Notes). Unfortunately, this was not done for the Housatonic. Because of this, it is difficult to differentiate between the effects on erosion rates of shear stress as compared with variable bulk properties.

RESPONSE 3-WL-5:

Please refer to the response to General Issue 7.

In a paper by Lick et al. (2005), approximate equations for sediment erosion rates are examined. It is shown that, for fine-grained, cohesive sediments, a valid formula is

$$E = 10^{-4}\left(\frac{\tau}{\tau_c}\right)^n \qquad (3)$$

where E is the erosion rate, τ is the shear stress, and τ_c is a critical shear stress defined as the shear stress at which an erosion rate of 10^{-4} cm/s occurs; τ_c depends on the particular sediment being tested and generally is a measured quantity. This equation is valid for fine-grained, cohesive sediments but **not** for coarse-grained, non-cohesive sediments.

For coarse-grained, non-cohesive sediments, the appropriate formula is

$$E = A(\tau - \tau_c)^n \qquad (4)$$

where A, τ_c, and n are functions of particle diameter but not a function of density. This equation is shown to be valid for coarse-grained, non-cohesive sediments but **not** for fine-grained, cohesive sediments.

To approximate erosion rates for all size sediments with a single, uniformly valid equation, the appropriate equation is

$$E = 10^{-4}\left(\frac{\tau - \tau_{cn}}{\tau_c - \tau_{cn}}\right)^n \qquad (5)$$

where $\tau_{cn}(d)$ is the critical shear stress for non-cohesive particles and is given by

$$\tau_{cn} = 0.414 \times 10^3 d \qquad (6)$$

where d is the particle diameter. Eq. (5) is uniformly valid for both cohesive and non-cohesive sediments. It reduces to Eq. (3) as $d \to 0$ and to Eq. (4) for large d.

In all the work we've done with Sedflume on the determination of erosion rates as a function of shear stress (the number of cores is on the order of 100), n in Eq. (5) is typically about 2 or more (see Lick et al., 2005 and Chapter 3 of Notes). Because of this, I suspect that the parameters n = 1.59 and n = 0.95 used in the Housatonic modeling (p. 4 of Attachment B.5) are incorrect.

RESPONSE 3-WL-6:

Please refer to the response to General Issue 7.

One reason for this may be (especially for n = 0.95) that Eq. (4) was used to describe erosion rates even though the sediments were fine-grained. In our work, when Eq. (4) was used to describe fine-grained sediments, the n determined by regression was quite low (1.31 in our experiments), but at the same time it was also shown that Eq. (4) was a poor approximation. However, Eq. (5) applied to the same data gave an $n \sim 2$ for all particle sizes, and the agreement between data and Eq. (5) was very good.

RESPONSE 3-WL-7:

Please refer to the response to General Issue 7.

Another reason for the low values of n determined for the Housatonic is that the above equations are only applicable to sediments which have the same bulk properties. In order to use these equations properly, sediments with similar bulk properties must be grouped together. Properties of sediments in a single sediment core generally vary with depth due to consolidation but also because of layering due to deposition after big events. Because of consolidation with depth, Sedflume measurements on one core will bias the value of n since cores at depth will be more consolidated, more difficult to erode, and will be measured later in the measurement cycle. Because the Density Profiler can give continuous density profiles as a function of depth with as little as 1 mm resolution, the use of the Density Profiler is important in determining the sediment structure and in interpreting sediment properties as a function of this structure.

RESPONSE 3-WL-8:

EPA agrees that the use of the University of California, Santa Barbara (UCSB) Density Profiler, if it had been applied to the Sedflume cores collected in the PSA, would have provided detailed bulk density profiles. However, EPA was not aware of the density profiler, and the USACE and Sandia National Laboratories contractors that performed the Sedflume analysis did not make this recommendation to EPA during the planning stages of the Sedflume analyses. In regard to the grouping of sediment with similar bulk densities for Sedflume analysis, please refer to the response to General Issue 7.

A way to group sediments with approximately the same bulk properties is as follows. (1) Separate sediments into fine, medium, and coarse sizes (this reduces the particle size effect) and (2) separate sediments by depth in the bottom sediments, e.g., 0 to 5 cm, 5 to 10 cm, 10 to 15 cm, etc. (the depth is a surrogate for changes in bulk density due to consolidation and hence this procedure normalizes the effect of density on the erosion rate).

RESPONSE 3-WL-9:

The approach for analyzing the data suggested above was followed, but it does not result in an improved characterization of the resuspension data. The results of the analysis are shown in Figures 3-WL-9a, b, and c. Each figure displays the data for a different particle size (D_{50} < 30 μm, 30 μm ≤ D_{50} < 100 μm, and D_{50} ≥ 100 μm, respectively) to reduce the particle size effect, and each panel presents data for different depth intervals (<5 cm, 5–10 cm, 10-15 cm, etc.) in an effort to account for changes in bulk density. When stratified in this manner, there is still considerable scatter in the data, an indication that factors other than particle size and bulk density are affecting the magnitude of the shear stress needed to resuspend Housatonic River sediment.

The values of n are crucial for extrapolating to and determining the effects of large storms (large stresses). The data as shown seemed to have large variability. Eq. (5) is probably a better

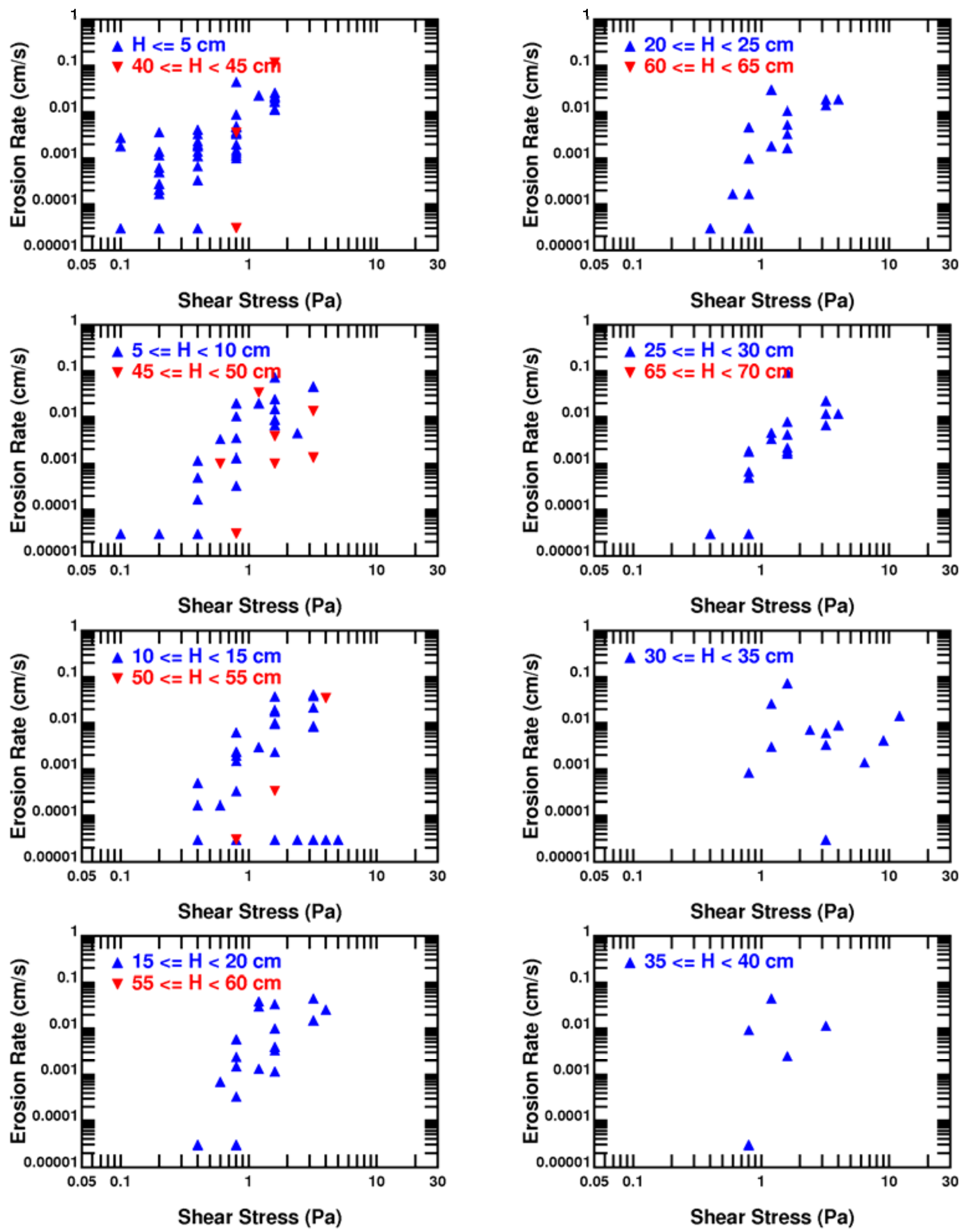


Figure 3-WL-9a Erosion Rate Versus Shear Stress (as a function of depth in core) for Sedflume Experiments on Sediment with Median Particle Diameter (D_{50}) Less than 30 µm

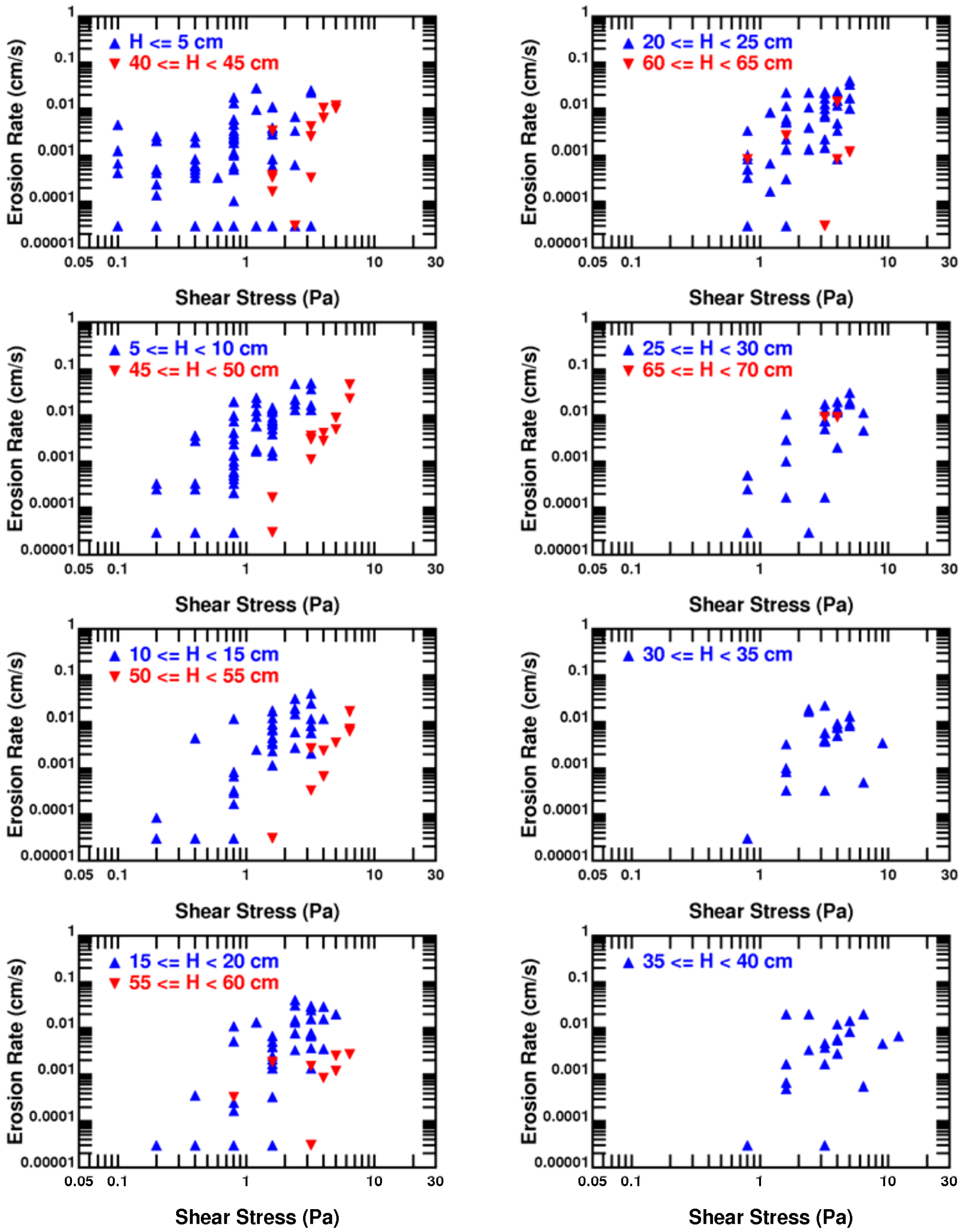


Figure 3-WL-9b Erosion Rate Versus Shear Stress (as a function of depth in core) for Sedflume Experiments on Sediment with Median Particle Diameter (D_{50}) Greater than or Equal to 30 µm and Less than 100 µm

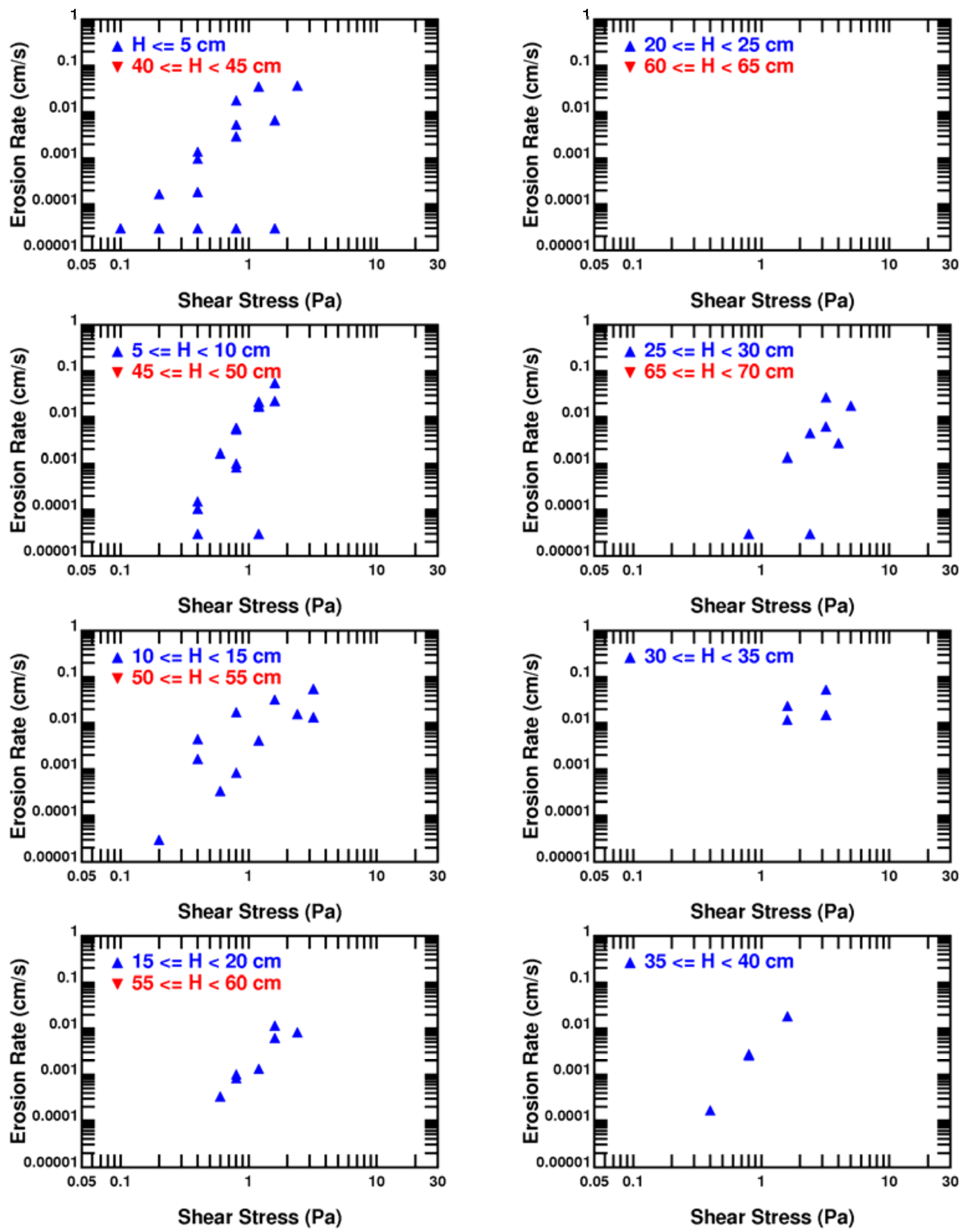


Figure 3-WL-9c Erosion Rate Versus Shear Stress (as a function of depth in core) for Sedflume Experiments on Sediment with Median Particle Diameter (D_{50}) Greater than or Equal to 100 µm

equation to approximate the data and may reduce this variability. In any event, the fact that $n \sim 2$ or more is a very strong experimental fact and hence must be considered seriously.

Whatever was done, Fig. 4-19 on p. 4-35 showing critical shear velocity as a function of grain size is incorrect. In all measurements that I am aware of (e.g., see Roberts et al., 1998; Lick et al., 2004, 2005; and Chapter 3 of Notes), the critical shear velocity (or critical shear stress) increases as grain size decreases beyond a minimum at about 100 to 200 μm.

RESPONSE 3-WL-10:

EPA reanalyzed the Sedflume data based on the Reviewer's comments; however, this reanalysis did not reduce the large variability in the data. Sensitivity analyses were also performed to evaluate the effect of larger exponents, but because of bed armoring effects the larger exponents did not produce the results anticipated by the Reviewer. The erosion of the non-cohesive fraction of the bed is described by the Garcia and Parker equations, and therefore, is not sensitive to changes in the value of n.

Bed armoring is an important process and causes large changes in bed shear stresses and hence large changes in erosion/deposition. This occurs, for example, when a layer of coarse sediments (as little as a few particle diameters thick) is deposited on a layer of finer, non-cohesive sediments. As the EPA model is presently configured, any deposited sediments are immediately mixed with the 6-inch surficial layer. Because of this, effective coarsening takes place very slowly (a small amount of added sediment has little effect on the average properties of the 6-inch layer). In reality, this mixing only occurs in a layer a few particle diameters thick, and this thin layer must be present in the model for realistic coarsening to occur (see SEDZLJ).

RESPONSE 3-WL-11:

The thickness of the surface layer used in the EFDC calibration was 7 cm, not 6 inches; however, EPA agrees with the Reviewer that the rate at which armoring of bed sediments occurs in the simulation is a function of the thickness of the surficial sediment layer, among other factors. Examination of the changes in bed sediment grain size distribution during the Phase 1 Calibration indicates that coarsening is occurring with the current 7-cm surface layer under flow conditions that would be expected to result in bed armoring. An example of bed coarsening was shown in Figure B.3-36 of the Model Calibration Report, which indicates that the process occurs relatively quickly under high-flow conditions.

As discussed in the response to General Issue 6, the thickness of the surface, or the biologically mixed, layer used in EFDC is being re-evaluated with respect to literature information on biological mixing of sediments in freshwater environments.

b. Flocculation and Deposition

Extensive work has been done on the flocculation of cohesive sediments including a simple model of time-dependent flocculation and measurements of settling speeds (Lick and Lick, 1988;

Burban et al. 1989, 1990; Lick et al., 1993; for a summary, see Chapter 4 of Notes). Experimental results show quite clearly that the steady-state median diameter of flocs is inversely proportional to CG (where C is sediment concentration and G is fluid shear) and inversely proportional to C when fluid shear is negligible. Since settling speeds are proportional to the floc diameter, this also demonstrates that settling velocities are also inversely proportional to CG and C. In general, flocculation is time-dependent as well as dependent on CG and C.

Eq. B3-34 on p. B.3-31, which is the formula for settling speed used in the modeling, has no dependence on time or fluid shear and has the incorrect dependence on sediment concentration. Fig. B.3-33 on p. 31 of B.3 Figures is completely inconsistent with experimental results.

RESPONSE 3-WL-12:

EPA notes that the relationships between the various parameters discussed by the Reviewer are not universally accepted by the scientific community. Indeed, other researchers have reported experimental results that reflect different relationships and invalidate the Reviewer's conclusions (Mehta and McAnally, 2002). The relationship between settling velocity and concentration used in the EFDC model is consistent with these other studies and also with the site-specific data.

The settling function presented in the Model Calibration Report (Eq. B.3-34) is a simple weighted average of settling velocities for low flow/low concentration washload and high flow/high concentration large aggregates. The Reviewer is correct that this formulation does not incorporate terms for time or fluid shear; however, EPA does not agree that the lack of dependence on time and/or fluid shear, or the equation used to express dependence on sediment concentration, is incorrect.

Initial calibration efforts were focused on using a settling velocity function for cohesive solids that depended on the product of concentration and shear stress. The resulting large longitudinal gradients in simulated depositional patterns were not consistent with patterns estimated from dated cores. In response to a recommendation provided during the Calibration Peer Review to use Dr. Lick's simple flocculation model (Section 4.4 of the Course Notes), EPA's modeling team implemented this flocculation model within EFDC. Although the model was able to represent the data from Dr. Lick's Couette flocculator, it was not able to reproduce the Housatonic River data; therefore, the flocculation model will not be used.

The settling function adopted for the Housatonic River model accounts for both the variations in floc sizes and settling speeds related to particle sources and rapid transport during storms, and localized conditions. During low-flow conditions (see Figure B.4-37 of the Model Calibration Report) a spatially constant low level of suspended solids was observed, which was modeled as a "washload" component with a low settling velocity. However, as exhibited during the high-flow event in October 2003, the relative concentrations of water column solids < 63 μm increased as total suspended solids concentrations increased (see Figures 4-13 through 4-17 of the MFD). The range of modeled floc settling speeds established during calibration is very close to the range of estimated settling speeds from the Phillips and Walling (2005) data.

Although the Reviewer does not believe that Figure B.3-33 is consistent with the experimental results he is familiar with, the relationship of settling velocity increasing with increasing concentration shown in the figure is consistent with the experimental data of several other researchers.

References:

Mehta, A.J. and W.H. McAnally. 2002. Fine-Grained Sediment Transport, Chapter 4. In *Sedimentation Engineering*, American Society of Civil Engineers Manual 54, Volume 2.

Phillips, J.M. and D.E. Walling. 2005. Intra-Storm and Seasonal Variations in the Effective Particle Size Characteristics and Effective Particle Density of Fluvial Suspended Sediment in the Exe Basin, Devon, United Kingdom. In I.G. Droppo, G.G. Leppard, S.N. Liss, T. Milligan, editors. *Flocculation in Natural and Engineered Environmental Systems*. CRC Press, Boca Raton, FL. pp 47-70.

Despite what EPA says in their response to my initial comments, I did not recommend or suggest a “flocculation formulation based on local instantaneous conditions”. That is EPA’s misinterpretation. On the other hand, EPA’s model of flocculation and settling speed, since it has no dependence on time, does depend only on local instantaneous conditions. As EPA states, that is incorrect.

RESPONSE 3-WL-13:

The formulation for settling velocity used in EFDC is not based on a model of flocculation. Because the EFDC formulation is based on the weighted average of washload and suspended load, it is appropriately calculated using local conditions. Testing in EFDC of the flocculation model recommended by the Reviewer did not reproduce the site-specific data.

For the correct determination of the flocculation and settling of cohesive sediments, a simple model of time-dependent flocculation is necessary and is available (Chapter 4 of Notes). This model reproduces all of our experimental results on the steady-state floc diameter as a function of sediment concentration and fluid shear and also the time-to-steady-state behavior as a function of time, fluid shear, and sediment concentration. Together with experimental results on settling speeds, this will give a correct and quantitative prediction of flocculation and settling.

The flocculation model is quite simple and is simply a single conservation equation for the average diameter of the flocs (which replaces the conservation equation for cohesive sediments already in the model) with a source term which determines the increase or decrease in the average size of the floc. The increase in computational time is negligible.

RESPONSE 3-WL-14:

Please refer to the response to General Issue 7.

c. PCB Flux and Depth of Mixing Layer in River

In the model as described (and in most water quality models), the depth of the mixing layer is absolutely crucial in the prediction of the long-term behavior of the PCB flux between the sediments and overlying water. In the absence of sediment resuspension/deposition, the time for natural recovery is directly proportional to the depth (thickness) of the mixing layer. Increasing the thickness by a factor of 2 increases the time for recovery by a factor of 2; decreasing the thickness by a factor of 2 decreases the time for recovery by a factor of 2, etc.

RESPONSE 3-WL-15:

EPA agrees that the thickness of the surface sediment layer is potentially an important parameter in long-term simulations of sediment PCB concentrations. Please refer also to the response to General Issue 6.

In the model, a six inch depth is prescribed. Why? This seems extraordinarily thick. GE says three inches would be better (thereby halving the time to natural recovery). Is there any theoretical or rational basis for 6 inches or 3 inches or any number? The depth of the mixing layer is a crucial parameter that determines the long-term behavior of natural recovery/remediation (basically it's the answer that you're looking for). Because of this, it should be a non-calibratable parameter.

RESPONSE 3-WL-16:

The question of mixing layer depth can be approached on both a theoretical and a rational basis. As part of the Phase 2 Calibration, EPA is evaluating hydrodynamic and biological information, as well as site-specific data, to constrain the range of plausible depths rather than allowing the value to range freely during calibration.

In fact, as explained below, I don't believe this parameter for this problem can be determined by calibration. As much as possible, this parameter should be based on scientific evidence and reasoning. Otherwise, as on the Fox and probably other locations, there will be different opinions and interminable arguments about the correct value for this parameter.

RESPONSE 3-WL-17:

Please refer to the response to General Issue 6.

To introduce some science into this argument, let me briefly review some information about PCB fluxes and the depth of the mixing layer.

The flux of PCBs between the sediments and overlying water occurs primarily due to sediment resuspension/deposition, molecular diffusion, and bioturbation. Pore-water convection and gas transport are generally less important but may be significant in some cases.

The flux of PCBs due to sediment resuspension/deposition is modeled **as a separate process** by calculating sediment resuspension/deposition and assuming equilibrium partitioning of the chemical. This modeling includes changes in sediment thicknesses due to resuspension/deposition.

RESPONSE 3-WL-18:

Please refer to the responses to General Issues 6, 7, and 8.

The flux due to molecular diffusion is generally ignored but it can be a significant process in itself and interacts with and modifies all the other processes. It is always present. Molecular diffusion of hydrophobic organic chemicals (HOCs) has been and is being investigated fairly thoroughly (Deane et al., 1999; Lick et al., 2004; Luo, 2005). For HOCs with large K_p's, it is known that (1) rapid changes in chemical concentration profiles are limited to a few millimeters near the sediment-water interface and (2) the magnitude of the flux is relatively large and constant for periods of 50 to 500 years and more. The profiles are limited to a few millimeters near the surface because in this region the diffusion is balanced by non-equilibrium sorption and only changes slowly with time. This leads to large gradients and hence high fluxes. Experiments give a mass transfer coefficient ($h \equiv q / C_o$) of approximately 10^{-5} cm/s or 1 cm/d. This is much higher than would be expected for a non-sorbing chemical. **This number is comparable to the 1.5 cm/d reported by the Housatonic modelers for their overall mass transfer coefficient.**

RESPONSE 3-WL-19:

Please refer to the response to General Issue 8.

Over time, the diffusion does penetrate into the sediments away from the interface, but at a very slow rate. As a first approximation, this rate can be shown to be h/K_p. For a $K_p = 10^5$ (approximately that for the Housatonic), this rate would be $10^{-5}/10^5$ or 10^{-10} cm/s = 10^{-5} cm/d = 4×10^{-3} cm/yr = 4 cm/1000 yr. It would take approximately 4000 years to diffuse through the 6 inch layer assumed in the modeling. Even for h = 1.5 cm/d as assumed in the model, it would take over 2000 years. If h = 1.5 cm/d and the mixing layer is 6 inches thick, this says that (except for resuspension/deposition) we can ignore natural recovery; it takes far too long.

RESPONSE 3-WL-20:

EPA agrees that the time necessary for natural recovery based solely on the rate of molecular diffusion would be considerable, but notes that molecular diffusion is only one of several processes that determine the rate of natural recovery, the most important of which are included in the modeling study.

A lot of work has been done on bioturbation, but the process is complicated because of many different organisms acting in different ways. As a result, bioturbation is not well understood quantitatively.

RESPONSE 3-WL-21:

EPA agrees that bioturbation is a complex process that is difficult to quantify. However, EPA has conducted a detailed literature review of freshwater bioturbation and evaluated the findings in the context of site-specific biological data to estimate the degree of biological mixing in the PSA. Please refer to the response to General Issue 6.

An important group of organisms is oligochaetes which burrow in the sediments. Their burrows generally extend 2 to 4 cm into the sediment; when disturbed, they may go deeper, up to as much as 10 cm, but only occasionally. They induce a contaminant flux by (1) diffusion of the chemical into and out of the burrow, (2) transporting fecal material to the sediment-water interface, and (3) slow mixing of the sediments due to their burrowing activities. Because of finite sorption rates, these processes are modified by molecular diffusion throughout but especially at the sediment-water interface, just as when molecular diffusion acts alone. The overall mass transfer coefficient is greater than that due to molecular diffusion alone by as much as a factor of 5 to 10, but only for the highest concentrations of oligochaetes as enumerated in the Great Lakes (EPA, 2004), and generally should be less (probably much less) than that. The magnitude of the mass transfer coefficient depends on the concentration of the organisms and decreases as the concentration decreases.

RESPONSE 3-WL-22:

EPA agrees that oligochaetes are important freshwater bioturbators and has considered the literature on these organisms, as well as others. Please refer also to the response to General Issue 6.

For other benthic organisms, the effects of bioturbation on the flux are probably smaller. Chironomids disturb the sediments to only 1 to 2 cm and the mass transfer coefficient seems to be smaller than that for oligochaetes. Hyallela cause even a smaller effect.

RESPONSE 3-WL-23:

Please refer to the response to General Issue 6.

In summary, benthic organisms disturb sediments primarily to a depth of 2 to 4 cm with some much smaller disturbances to as much as 10 cm. The mass transfer coefficient is on the order of 1 to 10 cm/d depending on the types and concentrations of organisms and is probably closer to 1 than it is to 10.

These numbers are approximate and subject to change, but they are the right order of magnitude. They suggest that the depth of a well-mixed layer due to molecular diffusion and bioturbation is on the order of 2 to 4 cm; some lesser disturbances may extend to as much as 10 cm. Sediment resuspension/deposition probably acts to depths greater than this; but this process is considered separately and should not be included when considering the depth of the mixing layer.

RESPONSE 3-WL-24:

Please refer to the response to General Issue 6.

EPA determined the magnitude of the PCB diffusive flux by assuming that, in a reach of the Housatonic, resuspension and deposition of sediments were negligible and therefore the entire flux was due to diffusion. This diffusion was reported as h = 1.5 cm/d; however, for PCBs, the rate of diffusion into sediments is h/Kp, or less than 0.1 mm/yr. In other words, this rate is equivalent to resuspension/deposition of about 0.1 mm/yr. Do we really know that resuspension/deposition is less than 0.1 mm/yr in this reach of the river? I doubt it.

RESPONSE 3-WL-25:

Please refer to the responses to General Issues 7 and 8.

Therefore, the assumption of h = 1.5 cm/d is questionable; this value is probably too low based on laboratory experiments. A better approximation is needed, preferably based on laboratory experiments and field measurements. This is another example where the parameter may be very difficult or even impossible to determine accurately by calibration.

RESPONSE 3-WL-26:

Please refer to the response to General Issue 8.

Neither pore-water convection nor gas transport has been shown to be significant in the Housatonic.

RESPONSE 3-WL-27:

The Reviewer is correct that data were not collected that could be used to describe pore-water convection or gas transport. PCB transport associated with groundwater influx is included in the model as part of a lumped parameter termed “vertical flux” that combines the processes of advection and diffusion. In the Conceptual Site Model (Section 4.5.2 of the Final Modeling Framework Design), the processes of advection and diffusion are discussed in some detail, and the lumped processes are concluded to be an important factor for the spatial distribution of water column PCBs. It is noted that advective flux is considered to be a calibration parameter, and synoptic bed sediment and water column data are available to calibrate the flux.

As stated above, if erosion/deposition is ignored and the contaminants have a Kp = 10^5, this indicates a natural recovery time of approximately 2000 years. On this basis, natural recovery is not an option; it takes too long. This applies throughout the river – bank to bank and from the confluence to Woods Pond. The entire river needs to be dredged and/or capped. No modeling is necessary for this conclusion.

RESPONSE 3-WL-28:

EPA does not dispute this conclusion based on the assumptions presented by the Reviewer; however, EPA notes that erosion and deposition are important processes that should not be, and have not been, ignored.

What happens when erosion/deposition is considered? Since the n's in EPA's erosion formulas are relatively small, I doubt that erosion/deposition will modify these results, even during big events, i.e., erosion will not penetrate down to the clean base sediments and there will not be enough erosion and hence subsequent deposition to cover the contaminated sediments by more than six inches of clean sediments in a reasonable time. These estimates should be checked by simple transport calculations (a big event and estimates of long-term deposition), but they are consistent with existing model results; my belief is that, to a first approximation, they are correct - based on EPA parameters.

RESPONSE 3-WL-29:

The Reviewer's concept that erosion through the entire depth of contaminated sediment is required to reach a supply of clean sediment that would be the only source contributing to natural attenuation ignores the solids entering the PSA from the East and West Branches, other tributaries, and runoff from the contiguous drainage area. PCB contamination extends as deep as 6 feet in Reaches 5A and 5B, 3 feet in Reach 5C, and 8.5 feet in Reach 6 (as reported in the Modeling Framework Design); therefore, the clean sediment is located deep in the sediment bed. Underlying "clean" sediment, in many cases, is a consolidated silty clay deposit, which is not easily erodable. The redistribution of contaminated sediment during extreme events will be evaluated and presented in the Model Validation Report. Additional discussion on the topic of resuspension and deposition is provided in the response to General Issue 7.

If dredging is done, sediments must be dredged down to clean base sediments, bank to bank, and along the entire river from the confluence to Woods Pond Dam. This follows from the model and EPA parameters.

If capping is done, the cap must be at least the mixing layer thickness plus whatever consideration of erosion requires.

RESPONSE 3-WL-30:

The evaluation of remedial alternatives will be performed in the Corrective Measures Study (CMS), and must consider a number of factors (beyond the model results) specified in the RCRA permit.

These conclusions follow from the EPA parameters and simple estimates – no lengthy calculations are needed. However, I don't really believe these conclusions. They may be correct, but the proof isn't there. The reasons that these conclusion may not be correct are (1) the assumed value of 6 inches for the thickness of the mixing layer is much too large and has no justification

RESPONSE 3-WL-31:

Please refer to the response to General Issue 6.

and (2) the assumed values of n lead to low erosion during big events and are inconsistent with experimental results. Both of these assumptions are inadequate.

RESPONSE 3-WL-32:

Please refer to the response to General Issue 7.

The use of only one grid cell across the river just exacerbates the problem.

RESPONSE 3-WL-33:

Please refer to the response to General Issue 4.

d. PCB Flux on the Floodplain

In the calculation of the PCB diffusive flux on the floodplain, since nothing else was said, the parameters such as depth of mixing layer and PCB mass transfer coefficient are the same as those in the river.

RESPONSE 3-WL-34:

The Reviewer is correct; the values for bed layering and the mass transfer coefficient used for the floodplain in the Phase 1 Calibration were the same as those applied in the river. Diffusive flux on the floodplain occurs only when the floodplain is inundated.

I assume there are minimal benthic organisms on the floodplain (even if present at any time, they wouldn't survive dry conditions), and therefore the PCB flux should be primarily due to molecular diffusion. For hydrophobic organic chemicals, such as highly chlorinated PCBs, the magnitude of molecular diffusion is comparable to but generally less than diffusion by benthic organisms; however, as described above, molecular diffusion behaves quite differently compared to bioturbation and certainly does not act over a mixing depth of 6 inches (Deane et al., 1998).

RESPONSE 3-WL-35:

The Reviewer is correct that benthic invertebrates, which are by definition aquatic organisms, do not inhabit the floodplain. However, as indicated by another Reviewer during the Peer Review, there are numerous terrestrial invertebrates, notably earthworms, that rework soils and whose activity would contribute to mixing of floodplain soil and associated PCBs. These organisms, although taxonomically different from those found in permanently wetted areas, contribute significantly to biological mixing of floodplain soil. According to Vrije University (2005), "one of the processes which may be important with respect to pollutants in soil is the presence of bioturbators, like earthworms. By mixing the soil, earthworms affect the distribution of pollutants in the soil and change physical soil properties, which have an influence on the availability of contaminants."

As part of the data collection in support of the Housatonic River Ecological Risk Assessment, EPA conducted soil invertebrate surveys at three locations in the floodplain in August 2000 and confirmed that several species of earthworms and other soil-reworking invertebrates are present in PSA floodplain soil. The earthworm and soil invertebrate sampling programs are discussed in Appendix A (Ecological Characterization) of the Ecological Risk Assessment. Please refer also to the response to General Issue 6.

Reference:

Vrije University. 2005. The effects of earthworm bioturbation on metal availability in river floodplains. Life Sciences – Animal Ecology Department. Faculteit der Aard – en Levenswetenschappen, Vrije Universiteit, Amsterdam, The Netherlands. URL: http://www.bio.vu.nl/do/Research/natureconservation/Earthworm_bioturbation

e. Grid Size and Number of Cells in River

With the present grid, the width of the river is generally approximated as one cell. In most rivers (and this includes the Housatonic; see cross sections in BBL report), there are large differences in erosion between the deeper parts and the shallower parts along a cross-section of a river. In particular, erosion/deposition in the deeper parts is not the same as the erosion/deposition averaged across the river.

Predicting the dissimilar amounts of erosion in the deeper parts and erosion/deposition in the shallower parts is crucial in predicting the long-term exposure of PCBs by erosion in the bottom sediments and/or natural recovery by deposition. Averaging across the channel does not describe the erosion/deposition processes accurately. A minimum of three cells across the river (two shallow, near-shore cells and one deeper, center cell) should be used.

RESPONSE 3-WL-36:

Please refer to the response to General Issue 4.

Quite extensive and high quality work has been done on the modeling. However, as described above, the model has serious deficiencies as far as the descriptions of (a) erosion, especially at high shear stresses, (b) flocculation and deposition, (c) PCB flux and depth of mixing layer in the river, and (d) PCB flux and depth of mixing layer on the floodplain. In addition, the spatial scale of the model in the river is inadequate to even approximately describe the variability of erosion/deposition across the river, and hence any remediation activities in the river. For these reasons, the model as is does not adequately account for the relevant processes affecting PCB fate, transport, and bioaccumulation in the Housatonic River. However, because of the extensive work already done and existing work done elsewhere, the model can be modified so that it can adequately predict PCB fate, transport, and bioaccumulation (see response to Question 6).

RESPONSE 3-WL-37:

Please refer to the responses to General Issues 4, 6, 7, and 8.

E. John List:

In short, the answer is no. The basic problem is with the specific application of EFDC model to the sediments in the Housatonic Valley. The difficulty arises in part because the PCB concentrations within the sediments show an extreme spatial variability. This variability exists on a very small length scale (see slides Number 3 and 4 of the presentation by Dick McGrath to the Peer Review Panel on April 13)

RESPONSE 3-JL-1:

Please refer to the response to General Issue 5.

and, what is even more important, it seems that this small-scale spatial variability is carried over essentially uniformly to the large scale, as is evident on Figure 5-26 of the calibration report. These figures show that PCB concentrations in the top six inches of sediment range almost uniformly from 0.5 mg/kg to 200 mg/kg (approximately three orders of magnitude) over 11 miles of the river valley.

RESPONSE 3-JL-2:

Please refer to the response to General Issue 5.

There is no explanation for this essentially uniform distribution of extreme variability in the calibration report. When Ed Garland (presumptive leader of the modeling team) was directly asked for an explanation for this variability at the Peer Review Panel Meeting he responded that he did not know, and nobody from the EPA consulting team volunteered an explanation, other than to state it was a commonly observed phenomenon with PCB contaminated sediments.

RESPONSE 3-JL-3:

Please refer to the response to General Issue 5.

In the absence of any explanation for why this variability is present it is difficult to believe that the modeling exercise, which deals only in spatial averages across the entire width of the basic river channel, can properly represent the fate and transport of the PCB.

RESPONSE 3-JL-4:

Please refer to the responses to General Issues 1 and 5.

The key point here is that there is no reason to believe that this extreme variability in the sediment concentration of PCB has not existed for many years; for it seems unreasonable to believe that it is a recent phenomenon. It is also unlikely that it is a mere sampling artifact. In other words, all of the sediment erosion and deposition since major PCB releases occurred has not caused the concentration of PCB's to average out, even over relatively short horizontal scales of a few meters. The importance of this observation is that the EFDC fate and transport model uses *spatially-averaged* PCB concentration data and therefore cannot possibly hope to reproduce either the observed current spatial variability or predict any future variability.

RESPONSE 3-JL-5:

Please refer to the responses to General Issues 1 and 5.

A long term application of the model is therefore simply going to smooth out the variance in concentration to produce a uniform concentration distribution; something that has not occurred naturally, at least so far.

RESPONSE 3-JL-6:

As discussed below (see Response 3-JL-7) and in the responses to General Issues 1 and 5, EPA does not agree that it is necessary for the model to reproduce the small-scale spatial variability in sediment PCB concentrations observed in the data.

In some respects, it is somewhat analogous to trying to predict the maximum force on a structure in a water body using a hydrodynamic model that includes tides but no waves. In this case a model is being used to project changes in average concentration over a relatively large spatial element when this large element includes sub-elements that have almost three orders of magnitude variability in concentration that are "washed out" in the averaging process. Since the output of the model provides concentration averages there can be no hope of it reproducing the observed spatial variation. In fact, since it is not known exactly what process sustains the spatial variability it is seems entirely possible that the EFDC model does not even have that transport process properly represented. It may well be that the spatial variation is a legacy of the manner in which PCB releases occurred in the past and will not occur in the future; in the absence of any explanation we simply do not know.

RESPONSE 3-JL-7:

As discussed in several other areas of this Responsiveness Summary, primarily in the response to General Issue 1, EPA believes that reproduction of small-scale spatial variability, or variability generally, is not necessary in the model output. Integration of modeling results over larger spatial and temporal scales is entirely appropriate because the primary goal of the modeling study is to predict average tissue concentrations in biotic receptors, primarily fish. The majority of these receptors integrate exposure concentrations over space and time and therefore, reflect the average of the contaminant concentrations to which they are exposed.

An additional goal is to examine the effectiveness of remedial alternatives, again on a scale far larger than the grid size used in the model. Therefore, reproduction of spatial variability has no practical value toward achieving the goals of the modeling study.

Furthermore, EPA does not agree that the analogy posed by the Reviewer is applicable to the modeling study. Unlike physical structures, which may well be affected by the maximum hydraulic force rather than the average, biota in the system act as integrators of the exposure concentrations; therefore, they reflect the average rather than the maximum.

There is no compelling reason to believe that the observed small-scale spatial variability in PCB concentrations in PSA sediment is the result of the manner in which PCBs were released from the GE facility. More likely, the variability reflects the integrated results of several processes that are already incorporated into the modeling framework, and some that may not be included. Regardless, because it is not necessary to reproduce the small-scale spatial variability in contaminant concentrations, it is not necessary for the model to explicitly include the process or processes that control the variability.

It is recognized that if, in the model, "average" particles are entrained by the stream, transported and dropped at some new location, then it should lead in time to the generation of a sediment and water column with average concentrations of PCB, which are then input to the FCM model.

RESPONSE 3-JL-8:

EPA agrees with the Reviewer's observation that the model appropriately evaluates the "average" particles resulting in concentrations of PCBs that are averages suitable for input to FCM. As discussed in the preceding response (3-JL-7), this is appropriate because biota integrate spatially and temporally variable exposure concentrations.

So, why not simply use the average concentrations and forget about the variability? The point is the use of spatially-averaged input data leads to a reduction in the variance of PCB concentrations that is solely an artifact of the modeling process, something that we have already seen to occur.

RESPONSE 3-JL-9:

Please refer to the response to General Issue 10.

On the other hand, the modeling process goes to great lengths to partition the PCB by sediment size while at the same time ignoring the spatial variability of the erosion/deposition processes across the river, where in fact there is significant particle size variability. Given the inherent non-linearity of the sediment transport process, it is not at all clear that the approach of taking a single modeling element across the river is going to lead to the correct results.

RESPONSE 3-JL-10:

Please refer to the response to General Issue 4.

Regardless of the reason for the PCB spatial concentration variations, this failure of the model to directly address the spatial variability in PCB concentration, which appears to be uniformly distributed along the river (see Figure 5-26), is a serious problem. It is highlighted by the fact that most of the calibration exercises performed for the EFDC model to date simply address the prediction of *averages* and the comparison of *averages* of field data and *averages* of model output (e.g., Figure B.4-46).

RESPONSE 3-JL-11:

As discussed in numerous responses to specific comments in this section of the Responsiveness Summary, and in the responses to General Issues 1 and 10, EPA does not agree that it is necessary for the model to reproduce the spatial variability in PCB concentrations observed in the data. The model result of primary interest is contaminant concentrations in biota, which integrate the variability in exposure concentrations.

The ability of a model to predict the *variance* in a distribution can be just as important as predicting the mean, especially where confidence limits on the predicted result, or the resolution of comparative remedial hypotheses, are important. There is nowhere (at least nowhere that I could find) in the PCB fate and transport calibration presentation that addresses the predicted *variance* in PCB concentration or flux (e.g., Figure 5-24) and calibrates this prediction with field data. It is implicitly presented in Figure 5-30 for tPCB water concentrations, but there is no formal comparison that I could find for the PCB fluxes, which are surely a very important part of the remediation modeling.

RESPONSE 3-JL-12:

Please refer to the response to General Issue 10.

One PCB transport mechanism that appears to have been neglected is the contribution from groundwater inflow into the stream. I did not see where the mass flux of PCB to the stream by this mechanism had been properly quantified. My own experience in measuring groundwater inflow to streams indicates that this contribution can be quite substantial, and at times can far surpass any diffusive flux out of the river bed sediments.

RESPONSE 3-JL-13:

Due to the hydrophobic behavior of PCBs generally, and particularly of the highly chlorinated mixture of congeners present at this site, transport of PCBs present in the sediment bed via groundwater advection is, by itself, a process of minor importance for the model. In addition, in much of the PSA the glacial Lake

Housatonic deposit, consisting of a fairly tight silty-clay layer observed at depth in the cores, likely controls the amount of groundwater advection that occurs directly to the stream channel. As discussed in Response 3-DE-5 above, groundwater inflow is included in the model as part of a lumped parameter termed “vertical flux” that combines the processes of advection and diffusion. In the Conceptual Site Model (Section 4.5.2 of the final Modeling Framework Design), the processes of advection and diffusion are discussed in some detail, and the lumped processes are concluded to be “. . . an important factor [for the] spatial distribution of water column PCBs. It was noted in that response that advective flux is considered to be a calibration parameter and synoptic bed sediment and water column data were available to calibrate the flux. . .”

Another mechanism for PCB loss that should be included in any analysis of the long term remediation process is the volatilization of PCB from the river flood plain sediments. Experience with other chlorinated hydrocarbons (DDT) has shown that the half-life for such compounds in soils exposed to solar heating can be in the range of 10-12 years.

RESPONSE 3-JL-14:

Please refer to the response to General Issue 5.

Methodologies for Evaluating Sensitivity and Uncertainty

4. Based on your technical judgment, have adequate methodologies been employed to evaluate the sensitivity of the model to descriptions of the relevant processes, and to evaluate uncertainties of model predictions?

E. Adams:

The analysis of sensitivity and uncertainty has been performed in the context of historical conditions; it should be repeated for future remedial conditions. That is, the model should be set up to assess the effectiveness of different remediation measures, and the sensitivity of this effectiveness to various model parameters assessed. This may lead to quite different conclusions. In some cases is might also be a lot easier. For example, if dredging is being considered, it is easy to simulate the effect in the model: the PCBs will simply be removed and the effectiveness of this remedial option (at least in a relative sense) will be totally insensitive to most model parameters. (There would be some short term disturbances during the actual dredging operation, but these would need to be simulated by different models, for which the current calibration study is irrelevant. Such disturbances are often considered part of an environmental impact assessment, but have not been mentioned; are they something to be considered?) Another example is capping which will introduce strong vertical gradients in PCB concentration within the sediments. Under this option the model may become much more sensitive to the way in which mixing between sediment layers is represented. Indeed, it may not be possible to calibrate some parameters based on historical field data (e.g., because the gradients are insufficient) so recourse must be made to literature values.

RESPONSE 4-EA-1:

EPA agrees that the sensitivity of the model to certain parameters may change in response to potential remedial alternatives and that sensitivity analyses performed in the context of historical conditions may not be representative of model sensitivities under future remedial conditions. This possibility will be considered when the model is used to simulate remediation scenarios.

A major difficulty with sensitivity and uncertainty analysis is that the model package is very expensive to run (owing principally to the long run time of EFDC reflecting, in turn, the small time step), and this problem will only get worse as the length of the simulation increases (from 14 months to order 20 years). Here I think the consultants could have tried to "think outside the box" a bit more with respect to the way the hydrodynamics, sediment transport and water quality were coupled. For example, the hydrodynamics could be run off-line and saved. And perhaps synthetic hydrologic sequences could be used. Or different grids could be used for the in-bank versus over-bank flows. QEA/GE made several suggestions along these lines at the Peer Review Meeting. If there is time, these should be seriously considered as a way to decrease computation time and allow the modeling team to afford to be able to run additional calculations.

RESPONSE 4-EA-2:

Please refer to the response to General Issue 3.

W. Frank Bohlen:

The methodologies used to evaluate sensitivity are generally accepted and adequate. Any deficiencies in the sensitivity analysis have more to do with omission. e.g. a test of the model sensitivity to upstream boundary conditions, noted above.

RESPONSE 4-FB-1:

The sensitivity analyses performed included upstream boundary inflows, TSS, and PCBs. These results are discussed in Sections B.2.4, B.3.4, and B.4.4, respectively, of the Model Calibration Report.

Douglas Endicott:

As mentioned previously, I thought the sensitivity analyses were well done. A couple of important parameters were omitted from the sensitivity analysis of EFDC, the settling rate and composition of particle types at the upstream boundary.

RESPONSE 4-DE-1:

The sensitivity analyses performed for EFDC included cohesive class settling velocity and critical shear stress for cohesive class deposition, and non-cohesive class particle diameters, which affect non-cohesive class settling velocities. These results are discussed in Section B.3.4 of the Model Calibration Report. The composition of particle types at the upstream boundary was not included as part of the sensitivity analyses, but will be considered as part of the Phase 2 Calibration effort.

Significantly, the sensitivity of the EPDC model predictions to the grid resolution was not included in the report.

RESPONSE 4-DE-2:

Grid resolution was addressed in the final MFD; therefore, it was not reevaluated as part of the model calibration. Please refer also to the response to General Issue 4.

The modeling team will need to revisit sensitivity at latter stages of the project, when longer-term simulations are run. At validation stage (for example), surficial sediment bed thickness should be explored in the sensitivity analysis.

RESPONSE 4-DE-3:

EPA agrees to revisit the sensitivity of the model predictions to the specification of the bed layering as part of the longer-term simulations (see Response 4-DE-1 above). EPA recognizes the importance of the specification of bed layer thickness in longer-term simulations, and plans to include it in the sensitivity analysis (see Section 7.3.4 of the Model Calibration Report).

Formal uncertainty analysis was only conducted on the food chain model. It may be the only one of the models for which uncertainty analysis can be practically accomplished.

RESPONSE 4-DE-4:

Please refer to the response to General Issue 11.

Any analysis of model uncertainty should address propagation of uncertainty between the models, and include uncertainty in statistical models (rating curves) predicting upstream boundary conditions, as well as uncertainty in flow measurements.

RESPONSE 4-DE-5:

Please refer to the response to General Issue 11.

I do not favor "bounding" analysis (as suggested by GE) as a shortcut to uncertainty. It is too easily to subjectively manipulate such an approach to produce a desired outcome. Bounding analysis should be used only if absolutely necessary due to computational constraints.

RESPONSE 4-DE-6:

Please refer to the response to General Issue 11.

Marcelo H. Garcia:

Sensitivity analyses of the PCB transport and fate model seems adequate. However, an evaluation of PCB fluxes between the river and the floodplain during overbank flow conditions is needed. Since stream bank erosion was not included in the calibration, it is not possible to know how sensitive the model predictions will be when such process is included.

RESPONSE 4-MG-1:

Please refer to the response to General Issue 12.

The uncertainty analysis has been done only for the Food Chain Model (FCM). An uncertainty analysis for the rest of the model components should be conducted, included. the uncertainty once all the models are linked together.

RESPONSE 4-MG-2:

Please refer to the response to General Issue 11.

Frank Gobas:

The methodologies used for the sensitivity analysis of the HSFP, EFDC and the Bioaccumulation models were carried out in an appropriate fashion. However, the sensitivity analyses would be more insightful if it they would focus on the key issues that the model needs to address. For example, the temporal response of PCB concentrations in water, sediment and biota to remediation efforts (and associated PCB loading reductions) is a key objective of the model. It would be helpful if the sensitivity analysis could report on the effect of various model parameters on the temporal response of the PCB concentrations in the area of concern. This is particularly important for the model parameters dealing with the sediment mixing (i.e. depth of active sediment layer, bed sediment mixing, resuspension and sedimentation) which have the largest effect on the time response of the PCB concentrations. However, it also important for other model parameters such as flow rates and lipid contents of fish. I therefore recommend that the sensitivity analysis is further developed. The current analysis provides useful information about which parameters are the most sensitive. The second phase of the sensitivity analysis can focus on these parameters and address how they affect key characteristics of the model, such as the temporal response of PCB concentrations in water, sediment and biota as a result of remediation options.

RESPONSE 4-FG-1:

The sensitivity of the models to parameters that may have a greater bearing on model results in long-term simulations is being considered. EPA recognizes that sensitivity analysis can and should be performed at various stages of model evaluation, and for this reason, the sensitivity of the model has been investigated at multiple stages, including Test Reach simulations, preliminary calibration, and Phase 1 Calibration. During Phase 2 Calibration and Validation, other parameters, such as depths of biologically mixed sediment, may exhibit greater sensitivity.

The methodology used to conduct the uncertainty analysis contains some significant limitations throughout the entire modeling effort. There are several issues. First, the basic modeling strategy relies heavily on model calibration. In model calibration, the observed data are used to parameterize the model. This produces a model where observed data cannot be used as an independent data set to test the performance of the model.

RESPONSE 4-FG-2:

EPA disagrees that the model for which formal uncertainty assessment has already been conducted (FCM) “relies heavily on model calibration.” The general calibration strategy for FCM is described in Section C.3.2 of the Model Calibration Report, and emphasized EPA’s reluctance to calibrate parameters in an unconstrained manner:

> "Calibration of the model required small adjustments to model parameters that were consistent with ranges of plausible values identified during the parameterization stage. In general, there was a strong preference to retain the 'best estimate' values identified in the parameterization stage. Although additional refinements of numerous model parameters would increase the statistical 'goodness of fit' of the model, such changes would not improve the predictive ability of the model (as measured by validation against independent data) if the changes were not justified based on the underlying science. Where possible, comprehensive reviews of parameter values were performed to develop distributions of values; in most cases the central tendency value was adopted for the calibration, and the variation and incertitude in the parameter was evaluated in the uncertainty assessment."
>
> Where calibration (i.e., fine-tuning of initial best estimates) was conducted, it was documented in the report, such that the quantitative effect of these modifications could be assessed by Reviewers. Because the model was not extensively calibrated to duplicate the calibration data, these data provide a reasonably independent test of model performance.

Comparing model predictions against independent data is probably one of the best and simplest ways to assess the uncertainty of the model. However, this method cannot be used to its fullest advantage in the model due to the reliance of calibration to make the models work.

RESPONSE 4-FG-3:

As described in Response 4-FG-2 (above), calibration was not always required to make the models work. The FCM uncertainty assessment evaluated the potential for bias (if any) in the model introduced by the limited calibration of the initial parameter estimates. Figures 29 through 93 of Attachment C.17 of the Model Calibration Report indicate that the central tendencies of PCB concentrations were similar pre- and post-calibration for most species and reaches.

EPA agrees that comparisons against independent data comprise a useful component of model uncertainty assessment. The FCM validation stage will include such independent comparisons by applying the model to data collected outside the calibration period. In addition, application of FCM to areas of the river downstream of the PSA will also assist in the model validation. Please refer also to the responses to General Issues 2 and 11.

Even without access to independent data (e.g. PCB concentration data in sediments and biota), there is still considerable merit to using differences between observed and simulated data as a measure of model uncertainty. I recommend that this is added to the modeling strategy given the limitations of the Monte Carlo Simulation technique that is the main method used to assess model uncertainty in this study. There are various statistical methods to do this such as mean squared error or calculating confidence limits of the model bias discussed above. The resulting uncertainty calculated should be treated with some caution as the uncertainty has a tendency to underestimate the actual uncertainty.

RESPONSE 4-FG-4:

The model framework already includes analysis of the difference between measured and simulated values as a part of model uncertainty and/or validation:

- Model validation will entail tests of the Phase 2 calibrated model against independent data sets (i.e., temporal extrapolation).
- The downstream modeling exercise will extend the model to additional reaches of the river with associated independent data sets (i.e., spatial extrapolation).
- Statistical summaries of model fit have already been provided on both a percentage and absolute basis in the Model Calibration Report.

Please refer also to the responses to General Issues 11 and 12.

The second issue relates to the application of Monte Carlo simulations, which was conducted in the bioaccumulation model. The application of Monte Carlo simulations to complex models like EFDC and the bioaccumulation model is difficult. There are two conditions that need to be met for the Monte Carlo simulations to be informative. One is that the model variables included in the Monte Carlo simulations are independent and not correlated. This was done for the bioaccumulation model but the report does not provide details on how this was done. This issue could therefore be expanded and perhaps improved upon in future work.

RESPONSE 4-FG-5:

EPA agrees that the application of Monte Carlo analysis to complex models with potentially correlated parameters is challenging (see Response 1-JL-7). The assumption of independence of inputs was addressed in FCM in several ways:

- The list of input parameters for which distributions were developed was constrained by identifying the parameters to which output was most sensitive. In this manner the potential for complex interactions among numerous distributions was reduced.
- The known interactions were accounted for, when possible, by specifying distributions to eliminate strong correlations. For example, the model is sensitive to the ratio between food assimilation efficiency and chemical assimilation efficiency. Therefore, rather than specify separate distributions for these parameters (which are strongly related), the uncertainty analysis modeled only the ratio between the two parameters. In this manner, the uncertainty analysis captured the uncertainty in the model but without introducing unnecessary autocorrelation in Monte Carlo inputs.

- Where inputs were correlated but could not be eliminated, professional judgment was used to specify the form of the functional interaction among inputs. For example, prey switching assumptions were made to specify the linkages among the dietary matrices. This ensured that inputs were logical (e.g., sum of preferences always equal to 100%) and that changes in feeding preferences represented realistic prey-switching behavior (based on the life histories of the modeled organisms).

EPA documented the rationales for the selection of inputs and distributions in Section C.4.2.1.2 of the Model Calibration Report. Please refer also to the response to General Issue 11.

A second condition for an informative Monte Carlo simulation analysis is that the variability and error in the model variables can be determined or are known. For some model variables this can be done relatively easily, while for others (e.g. feeding preferences, growth rates) this is very difficult. The report does provide information on this issue and the authors are doing a good job to deal with this difficult issue.

RESPONSE 4-FG-6:

EPA agrees that the specification of input distributions was challenging for some parameters. Some parameters could be specified using routine statistical analysis of multiple data points; others, however, required professional judgment to evaluate multiple lines of evidence and/or to incorporate non-quantitative information. Please refer also to the response to General Issue 11.

An issue that requires further investigation, in my view, is why the MCS calculated uncertainty in the concentrations of PCBs in biota is considerably less than the observed uncertainty. The latter is indeed not impossible as the MCS method does not capture all sources of uncertainty. However, it raises issues about the value of the uncertainty analysis and how the results of the MC simulations should be interpreted when applying the model.

RESPONSE 4-FG-7:

The reason that the Monte Carlo distributions are narrower than the distributions derived from field data is that the Monte Carlo analysis generally did not incorporate all sources of variability in individual fish. Uncertainty consists of both variability and model incertitude. The Monte Carlo analysis attempted to characterize the incertitude associated with mean concentrations of fish for each combination of species, age, and reach. The distributions could easily have been made wider by specifying distributions around additional parameters that are relevant to individual fish PCB concentrations, such as variability in exposure to media at the base of the food web (see Figure 1-JL-7 for sediment concentration example). However, because the primary objective of the model was to simulate mean concentrations instead of individual concentrations (see response to General Issue 1), these additional distributions were not specified. Please refer also to the response to General Issue 11, and to Responses O-DE-9 and 1-JL-7.

In terms of uncertainty analysis for the EFDC and FCM models, there is considerable room for improvement and additional work. For example, there is no uncertainty analysis for the EFDC model at this point.

RESPONSE 4-FG-8:

Please refer to the response to General Issue 11.

As for the FCM model, the MC simulations provide useful information but the report makes qualitative statements (i.e. "two-fold" in an number of places e.g. p.C.4-24), which do not appear to be representative of the real model uncertainty as demonstrated by differences between model predicted and empirical PCB concentrations in biota.

RESPONSE 4-FG-9:

The qualitative statements regarding "two-fold" differences are accurate when applied to comparisons between mean observed tissue concentrations and FCM-simulated concentrations. Such comparisons are represented by yellow square symbols in Figure C.3-51. For sample sizes greater than 6, all of the tPCB comparisons were within a factor of 2 (represented by a thick dashed line), and the majority of individual congener comparisons were also within a factor of 2, which was established in the Modeling Study QAPP and the final MFD as the target tolerance for calibration of the bioaccumulation model (see final MFD, Table I-1 in Appendix I). The figure indicates that increased sample size results in stronger agreement between measured and simulated data, which is to be expected given the large variability in individual PCB concentrations.

In Figure C.3-51, EPA presented comparisons of measured and simulated data even if sample sizes were insufficient to calculate meaningful central tendency statistics (e.g., n = 1 cases represented by blue circle symbols). However, individual fish concentrations that deviate from simulated average values by more than a factor of 2 should not be interpreted as indications of lack of model fit, particularly because the individual tissue concentration variations in the field are more than an order of magnitude. The individual data points in Figure C.3-51 were plotted to assess the model bias (or lack thereof), not to test the MFD/QAPP performance criteria that apply only to central tendency concentrations. Please refer also to the response to General Issue 11, and to Responses O-DE-9 and 1-JL-7.

My recommendation is:

1. To include an assessment of model uncertainty based on a comparison of observed and predicted concentrations (The 95% confidence intervals of the model bias, discussed earlier, can be a useful tool to do this).

RESPONSE 4-FG-10:

Please refer to the response to General Issue 11 and to Response 2-FG-8.

2. To conduct the planned MC analyses considering the importance of conducting the analyses with non-correlated state variables and supporting the distributions of state variables with scientific data or appropriate and documented judgment.

RESPONSE 4-FG-11:

EPA has already considered intercorrelations among inputs. Please refer to the response to General Issue 11 and to Response 4-FG-5 above.

Finally, it is important to stress that both approaches have their pros and cons and that they only arrive at estimates of model uncertainty. In the application of the model this should be recognized.

RESPONSE 4-FG-12:

Please refer to the response to General Issue 11.

Wilbert Lick:

As described in the introductory material for Question 3, rates of erosion and deposition are both questionable. The fact that the model predicts suspended sediment concentrations fairly accurately does not imply that the rates of erosion and deposition are correct.

RESPONSE 4-WL-1:

Please refer to the response to General Issue 7.

The well-mixed layer is very thick. Because of this, its thickness is insensitive to diffusional flux processes and resistant to calibration. I think it will be almost impossible to determine by calibration; some scientific reasoning is necessary.

RESPONSE 4-WL-2:

Please refer to the response to General Issue 6 and to Response 3-WL-16.

Sensitivity and uncertainty analyses are meant to measure differences between the model calculations and observations and determine the best parameters for the model as specified. These tests do not determine whether the physical formulations/equations in the model are adequate or correct. The model as is has the incorrect physical formulations; and sensitivity/uncertainty analyses will not demonstrate this.

RESPONSE 4-WL-3:

EPA disagrees with the Reviewer's comment that the model "has the incorrect physical formulations"; additional discussion relative to this comment is provided in the response to General Issue 7. EPA also disagrees with the Reviewer's comment that one of the objectives of sensitivity and uncertainty analyses is to "determine the best parameters for the model…." This objective is part of the model calibration effort. Additional discussion of the uncertainty analyses planned for the Phase 2 Calibration is provided in the response to General Issue 11.

E. John List:

The sensitivities of model predictions to the relevant processes that have been included within the models seem to have been exhaustively evaluated. The issue is the uncertainties in the model predictions, which are of two types. One type is the inherent uncertainty associated with the uncertainty in the input data, which of course arises from the precision and accuracy of sample laboratory work and the inability of a limited number of samples to exactly define the statistics of a population. All of these uncertainties are quantifiable, although this has not been done to the level one would have liked to have seen.

RESPONSE 4-JL-1:

Please refer to the responses to General Issues 5 and 11.

For example, the uncertainty in the predictions of PCB and sediment stream fluxes appears to have been omitted (e.g., in Figure 5-24).

RESPONSE 4-JL-2:

As discussed in Section B.5.3 of the Model Calibration Report, and in response to a preliminary question received from the Panel, EPA did not believe that performing the uncertainty analysis over the relatively short calibration period would have been very informative for either HSPF or EFDC.

In addition, there has been no known attempt to conduct a rigorous uncertainty analysis at other sites using a model such as EFDC. Given this, and also recognizing the amount of resources required to perform the analyses for both EFDC and HSPF, EPA had hoped to receive input from the Panel on the process to be used to conduct the uncertainty analysis prior to expending the significant time and resources required to conduct such analyses. The uncertainty analysis performed for FCM during model calibration was very straightforward and required relatively few resources; a similar approach will be applied during model validation.

Another omission that could have easily been included is whether the mean PCB concentrations in the upper 6 inches of sediment show any statistically significant difference along the 11 mile river reach. This is exemplified by the model input data in Figure 5-26. Because there are no confidence limits on the estimates of the mean concentration of tPCB along the river reach it is not clear that the variations shown in the tPCB profile along the river have any statistical significance.

RESPONSE 4-JL-3:

Sediment PCB concentrations in the various modeling reaches were analyzed and discussed at some length in Section 4.5 of the RFI report (BBL and QEA, 2003). Although statistical hypothesis testing was not conducted, error bars (± 2 standard errors, approximately equal to a 95% confidence interval) are provided for the mean PCB concentration by PSA reach in Figure 4-8 of the RFI. Examination of the means and their confidence intervals suggests considerable overlap in surficial sediment PCB concentrations in the PSA, with the exception of Reach 5B, which appears to have generally lower concentrations than the other reaches.

Because the distributions of the sediment PCB concentration data in each PSA reach are strongly positively skewed (i.e., lognormal), the arithmetic means are not the best measures of central tendency for use in statistical hypothesis testing and statistical comparisons based on these means are not valid. To investigate whether there are significant differences among the four main channel reaches in the PSA (5A, 5B, 5C, and 6), two different statistical tests were used.

The null hypothesis tested was of the form:

H_0: $x_{5A} = x_{5B} = x_{5C} = x_6$,

where:

x = A measure of central tendency.

5A, 5B, 5C, and 6 = The subreaches used for FCM.

The testing was conducted on all surficial (0 to 15-cm) sediment tPCB data in the project database; non-detects were included at the detection limit, but there were very few non-detects in the data set.

Median Test

First, the median values for the four reaches (5A = 11.3, 5B = 4.42, 5C = 6.81, 6 = 17.8) were tested using the *median test* of Mood (1950, as discussed in Zar, 1999). This procedure involves the use of the X^2 distribution to test the number of data points in each reach that fall above or below the median value for all reaches combined. The results of this test were very highly significant ($p << 0.001$), and the null hypothesis was rejected (i.e., at least one reach mean was significantly different from the others). Although it is theoretically possible to conduct a posteriori multiple comparison testing of medians to determine which reaches are significantly different from the others, the test is not robust and requires equal sample sizes, which is not the case for the sediment PCB concentration data.

Analysis of Variance (ANOVA)

PCB concentrations in the four subreaches were also tested using one-way analysis of variance (ANOVA). The four groups of data were plotted and determined visually to be approximately lognormal, thus, the data were log-transformed ($\log_{10}$ x+1) and replotted. Based on a visual inspection, the transformation was successful in making all distributions approximately normal. The ANOVA based on the four transformed means (5A = 1.10, 5B = 0.73, 5C = 0.95, 6 = 1.21, equivalent to geometric means of 11.6, 4.37, 7.91, and 15.2, respectively) was very highly significant ($p << 0.001$), indicating that at least one of the means was significantly different from the others.

The differences among the reach means were tested pairwise using Tukey's honestly significant difference (HSD) test (Zar, 1999). All reach means were found to be significantly different from each other at $\alpha = 0.05$, although the difference between Reaches 5A and 6 was only marginally significant.

These results, which in part reflect the large amount of data available for the PSA, indicate that there are statistically significant differences between PCB concentrations in the surficial sediment of the modeling reaches. The differences, however, do not reflect a monotonic change with longitudinal distance downstream, in part because of differences in hydraulic regime, which in turn influence patterns of erosion and deposition, resulting in differences in grain size, TOC, and other parameters among the four reaches.

Reference:

Zar, J.H. 1999. *Biostatistical Analysis*. 4th Edition. Prentice Hall, Upper Saddle River, NJ. pp. 663.

The second type of uncertainty is created by the inability of a model to exactly replicate the processes active in the field and the very distinct possibility that a model may have inherent, and unrecognized, flaws. This type of uncertainty is very difficult to quantify and can only be properly defined by a careful calibration and validation procedure. If the calibration and validation procedures indicate, over a large number of evaluations, that the model is reproducing field results, then some degree of confidence can be ascribed to the output of the model.

RESPONSE 4-JL-4:

Please refer to the response to General Issue 11.

Of particular concern here is that despite reproducing the mean of several measured parameters adequately, the model is not indicating the proper degree of variance in the output, as was discussed above in the response to Charge Question No. 1.

RESPONSE 4-JL-5:

Please refer to the response to General Issue 10.

On the other hand, as pointed out by another member of the Peer Review Committee (Dr. Lick), the model may be getting the "right" results despite incorrect science, simply because there are two compensating basic errors.

RESPONSE 4-JL-6:

The basis for Dr. Lick's comment that the model may be getting the "right" results for the "wrong" reason was that the concentration of suspended solids could be fit with a wide range of combinations of settling and resuspension rates (see response to General Issue 7). Dr. Lick's assumption ignores the examination of the gradient between sorbed PCB concentrations in the water column and the sediment. Although different combinations of settling and resuspension rates could be assigned to produce the same solids flux between the sediment and water, the PCB flux from the sediment to the water column will increase with increasing resuspension (even if settling is increased) because the PCB concentrations on sediment solids are higher than on water column solids. Thus, the PCB concentrations simulated by the model provide an important check on the combination of settling and resuspension rates. Please refer to Response 1-DE-8 and the response to General Issue 7 for additional discussion.

Another issue in this regard was also commented upon above. The fact that the sediments show a very wide variability in concentration of PCB, three orders of magnitude on horizontal spatial scales of 2 meters or so and uniformly over the 11-mile river reach, is disturbing for two reasons. One is the fact that there is no known explanation for this and therefore the appropriate fate and transport mechanism cannot be included in the model. And the second is that despite the fact that the mechanism is not included in the model (because it is unknown), there is no proof that its omission is of no consequence to the outcome of the modeling exercise.

RESPONSE 4-JL-7:

Please refer to the response to General Issue 5.

Use of the Model to Evaluate Differences in Remedial Options

5. Is the uncertainty indicated by model-data differences sufficiently inconsequential to permit use of the model to predict differences in remedial options?

E. Adams:

See response to Question 4.

W. Frank Bohlen:

For the reasons discussed in Question 3 above answering this question seems premature.

Douglas Endicott:

This really is a judgment call that depends upon how ambitious the remedial options are. Uncertainty is "inconsequential" if it does not obscure the discrimination between outcomes (i.e., PCB body burdens in fish) of different scenarios.

RESPONSE 5-DE-1:

Please refer to the responses to General Issues 1 and 11.

At this point, uncertainty of EFDC predictions have not been determined. The uncertainty of food chain model predictions of total PCB bioaccumulation appear to be acceptable to evaluate remedial options. For most fish, the ratio of predicted 90^{th}-percentile to 10^{th}-percentile PCB concentrations is a factor of 3 to 5. This matches my expectations from other systems, and indicates a high-quality model.

RESPONSE 5-DE-2:

EPA agrees that the model performance compares favorably with results from modeling other systems. The uncertainty bounds (percentiles) likely overstate the uncertainty about central tendencies of PCB concentration because some input distributions in the Monte Carlo analysis included variability among individuals in addition to model incertitude. Further information on model performance will be available following model validation and extrapolation of the model downstream of the PSA.

Of course, this is just uncertainty due to model parameters, which is only a part of real model uncertainty. I am told that the really fatal model uncertainties are the things you don't know about and cannot be anticipated. That is why (aside from curiosity) I would like to see EPA collect more data on this system, as identified above (response to question 3).

RESPONSE 5-DE-3:

EPA believes the data already collected on this system are sufficient to achieve the goals of the modeling study. The large amount of data, combined with the reasonable results from the short-term calibration and the demonstrated risks at the site, suggest that the goals of the modeling study will be met with the data that have been collected and that the project should move forward.

Marcelo H. Garcia:

At this stage, the model has been used to predict sediment and PCB transport and fate during storm events. So the model-data differences might not be very important for the time and space scales considered for model calibration, but could very well increase in relevance for the scales needed to compare remedial actions.

RESPONSE 5-MG-1:

Please refer to the response to General Issue 2.

Frank Gobas:

As for the HSPF model, the model-data differences are sufficiently small to use the model to predict mean flow, TSS and temperature in the River following remedial actions.

RESPONSE 5-FG-1:

EPA agrees with the Reviewer's assessment.

On balance, the hydrodynamic model and sediment transport model appear to produce relatively small differences between model predictions and observations. The small differences between observations and predictions are partly caused by the calibration methodology which uses the observed data to make the model predictions. Hence, a good agreement between observations and predictions should be expected.

RESPONSE 5-FG-2:

Please refer to the response to General Issue 2.

It is unclear from the study so far how predictive the model really is and hence what the model's uncertainty is. This can de determined by conducting a model validation (better is model performance evaluation), where the observed data are not used to make the model predictions.

RESPONSE 5-FG-3:

Please refer to the response to General Issue 2.

However, despite some limitations in the approach so far, the model is a reasonable tool to start making certain predictions among remedial options.

The performance of the PCB fate model is only tested in its ability to estimate PCB mass in the water column. While the performance of the model as characterized by differences between observed and predicted concentrations are reasonable they do not shed much light on the ability of the model to estimate spatial differences in PCB concentrations (as concentrations of PCBs do not appear to show statistical differences among the stretches of the River of concern), or the model's ability to estimate the temporal response of the PCB concentrations in water and sediments in the River.

RESPONSE 5-FG-4:

Please refer to the response to General Issue 12.

If it is further considered that the model may not have fully represented some key fate processes, I recommend considerable caution in the application of the model in a predictive sense, in particular if the long term temporal response of the model is important. I think uncertainty analyses need to be added.

RESPONSE 5-FG-5:

Please refer to the response to General Issue 11.

Model – data differences in PCB concentrations in the bioaccumulation model are considerable despite significant calibration efforts.

RESPONSE 5-FG-6:

Please refer to Responses 1-JL-6 and 1-JL-7 (regarding model-data differences) and 4-FG-2 (regarding degree of calibration) above.

Biological data often exhibit a large degree of variability. Hence, it is not uncommon in bioaccumulation modeling efforts that there are significant discrepancies between predicted and observed concentrations. The latter should not be viewed as criticism of the model or an impediment in the application of the model. As long as the uncertainty in the model calculations is appropriately recognized, the results of the model can be interpreted accordingly and the model can be used productively to assess the impact of remedial actions.

RESPONSE 5-FG-7:

EPA generally agrees with the above comment. However, the term "discrepancy" should only be used for model-data comparisons that represent "apples-to-apples" conditions (i.e., central tendency of measured concentrations versus central tendency of simulated concentrations).

The issue of uncertainty requires further attention in the development of the bioaccumulation model. Currently, the report contains statements about "the majority of PCB tissue concentrations being within a factor of 2 of the deterministic values". These statements does not appear to be representative of the real model-data differences shown in Figure C.3-27 or even Figure C.3-28 (which I presume are mean concentration values). Only, when uncertainty is appropriately recognized, application of the model should be considered.

RESPONSE 5-FG-8:

Please refer to the response to General Issue 11 and to Response 4-FG-9 above.

Although I do not think that the uncertainty of the FCM model is correctly represented in the report, I do think that when this is done, the model can be used to predict differences in PCB concentrations in biota resulting from remedial options despite the fact that differences in observed and predicted concentrations are considerable.

RESPONSE 5-FG-9:

EPA disagrees with the conclusion that "differences in observed and predicted concentrations are considerable" when data and model simulations are considered as central tendencies. Please refer also to Responses 1-JL-6 and 1-JL-7.

It is possible that, in some cases, there may not be statistically significant differences in PCB concentrations in fish resulting from different remediation scenarios but, if so, this is important information to know.

RESPONSE 5-FG-10:

In evaluating different remedial options, both the magnitude of differences (effect size) and the significance of differences are important. Statistical significance measures are useful but can be misleading unless the statistical power of the tests and the costs (consequences) of both Type I and Type II statistical errors are considered. Please refer also to the response to General Issue 1.

Wilbert Lick:

No (see response to question 4).

RESPONSE 5-WL-1:

EPA notes that the uncertainty analyses for HSPF and EFDC were not performed as part of the model calibration and therefore, were not included in the Model Calibration Report. These analyses will be performed on the longer-term simulations and will be included in the Model Validation Report.

E. John List:

In order to resolve a formulated hypothesis that one remedial option has a better "performance" than another, (however that performance may be quantified), there has to be a comparative statistical test developed. The test must be capable of resolving the probability of both false positives and false negatives and knowledge of sample means alone is insufficient information to resolve any formulated hypothesis.

> **RESPONSE 5-JL-1:**
>
> EPA agrees that means, or other measures of central tendency, alone do not provide sufficient information for statistical tests; however, the model, as is the case with most models, does not output directly the information on variability needed to conduct such tests. The proposed procedure for evaluation of uncertainty of model output will provide estimates of model output variability that will allow some approximation of hypothesis-testing procedures. Evaluation of alternatives via more qualitative inspection of model results, combined with the output of the uncertainty analysis noted above, will also be used to provide useful information for decisionmakers.

There are certainly non-parametric tests that can be used, but these require some knowledge as to the nature of the distributions and the ability to resolve false negatives may be impaired.

> **RESPONSE 5-JL-2:**
>
> The usual purpose of non-parametric tests, also known as "distribution-free" procedures, versus parametric tests is to allow hypothesis testing without needing to specify the distribution of the data being tested. Non-parametric tests vary widely in their power and robustness, and can be more powerful than parametric tests if the assumptions for the equivalent parametric test are not met.

As previously discussed, there is definite uncertainty in the model predictions and some of this has not been quantified, which will definitely impair the ability to resolve statistical hypotheses regarding the relative performance of remedial options.

> **RESPONSE 5-JL-3:**
>
> Please refer to the response to General Issue 11.

More attention should be directed at developing the data required for hypothesis testing.

> **RESPONSE 5-JL-4:**
>
> The data collection phase of the project has concluded, and EPA believes that the data collected to support the modeling study are sufficient to support achievement of the goals of the modeling study. EPA also believes that the recently completed human health and ecological risk assessments demonstrate elevated risks at this site, and that it is prudent to proceed with the remainder of

the modeling study. Accordingly, with very limited exceptions, no additional data will be collected.

As discussed in several responses in this document, and particularly in the response to General Issue 11, EPA is investigating a number of methods to evaluate model uncertainty and believes that quantification and evaluation of model uncertainty will provide a better means of evaluating the differences between remedial alternatives than will a definitive statistical hypothesis-testing protocol.

Overall Adequacy of Calibration

6. Are the processes in the model calibrated to the extent necessary for predicting future conditions including future concentrations of PCBs in the environment under natural processes and under potential remedial options for sediments and floodplain soils in the Housatonic River in the reach below the confluence? If not, what additional work needs to be done to calibrate the model?

E. Adams:

See response to Questions 1, 2 and 4 which also include suggestions for improving model calibration.

W. Frank Bohlen:

No. The fundamental problem (beyond the issues discussed above) with the current model calibration is its limited duration. Since data seem to be available for the extension of this period it is my recommendation that the period be extended in year steps out to five years. As suggested above, the five year period appears to be sufficient to allow measureable change in sedimentation and the associated sediment/contaminant concentrations to occur.

RESPONSE 6-FB-1:

Please refer to the response to General Issue 2.

Douglas Endicott:

As I mentioned above (see response to question 1), we can only anticipate the calibration of processes that are influential to long-term model predictions. In this context, I am primarily concerned with the calibration of resuspension fluxes, the sediment-water diffusion flux, and the surficial sediment residence time as defined by the mixed layer thickness.

RESPONSE 6-DE-1:

Please refer to the responses to General Issues 6, 7, and 8.

I also worry about unquantified PCB sources that will remain after remediation is completed.

RESPONSE 6-DE-2:

Under the terms of the Consent Decree, there are requirements to address the sources of PCBs at the GE facility and other areas specified in the Consent Decree. In addition, MDEP is working with GE to address all known sources other than those specified in the Consent Decree. With regard to potential contributions from tributaries and the Pittsfield WWTP, please refer to Responses O-DE-8 and 3-DE-4.

There have been protracted arguments about whether the thickness of the surficial mixed sediment layer, or (alternatively) the bioavailable sediment layer, should be 6 inches as opposed to 3 or 4 inches. I think it is fair to say that each side prefers a number intended to produce a modeling outcome favorable to their own interest. EPA has not convincingly demonstrated that the current model parameterization of mixed layer depth is scientifically defensible. The truth is we have almost no system-specific data to guide the specification of this parameter, and there is little guidance currently available in the literature from other sites. All lines of evidence should be used to evaluate this parameter. That should include physical, chemical and biological data as well as what the model can tell us via calibration. The modeling team should consult with scientists familiar with the various biotic and abiotic benthic processes.

RESPONSE 6-DE-3:

Please refer to the response to General Issue 6. Additional specific responses to comments made above include:

- EPA does not have a preference for any specific value of the parameter other than to optimize the predictive value of the model.
- EPA demonstrated that application of a thick layer of biological mixing was an adequate simplifying assumption for the calibration period. The best “scientifically defensible” thickness for use in longer-term simulations was not determined at the time of the Model Calibration Report.

Site-specific data are available and these were considered in the specification of biological mixing depth (see the response to General Issue 6). EPA agrees that a weight-of-evidence approach is useful for specification of this parameter.

Finally, it should be recognized that simulations of long-term remediation alternatives will dramatically alter the sources and pathways followed by PCBs in the PSA. Instead of the dominant role played by advection of PCBs from the upstream boundary (the major flux pathway during the calibration period), PCB flux via diffusive and particulate fluxes from the sediment will eventually predominate. This will change the general sensitivity and uncertainty behavior of at least EFDC and may reveal errors that are not currently evident. I think some scenarios of this sort should be tested during model calibration, to ensure that the models behave in a way that is at least consistent with our overall understanding of the system. Models sometimes behave unusually when the major contaminant pathways are altered.

RESPONSE 6-DE-4:

Please refer to the response to General Issue 12.

Marcelo H. Garcia:

The calibrated model ability to predict future concentrations of PCBs as well as the impact of potential remedial actions in the Housatonic River and its floodplains, will be very limited. The main issue with the calibrated model is the size of the computational grid employed both for the river and its floodplain as well as the fact that the model has been calibrated for a relatively short period of field observations. While the latter is driven by data availability, the former can be solved with today's computational resources.

RESPONSE 6-MG-1:

Please refer to the responses to General Issues 2, 3, and 4.

My recommendation is to include stream bank erosion in the model since this is a very important process which has the potential for greatly impacting the modeling outcome. While the mechanics of stream bank erosion is not well understood, it is possible to assess how much sediment enters the channel though bank erosion with a simple 1-D model for meandering streams. Such model is described in an attached paper by Abad and Garcia (2005).

RESPONSE 6-MG-2:

Please refer to the response to General Issue 9.

There is a clear need for the model to be able to resolve flow and sediment transport within the channel and along its stream banks. To model mass transfer processes at the sediment-water interface fine-scale hydrodynamic (i.e. flow velocity, shear stresses) predictions are needed. As it stands, the calibrated model cannot do this. A 3D hydrodynamic and sediment transport model should be used for in-channel flows and a 2D model (like the one currently employed) for the floodplain. This is computationally challenging but can be done as shown in the attached paper by Rodriguez et al. (2004).

RESPONSE 6-MG-3:

EPA does not believe it would be advisable to use the approach suggested by the Reviewer. As stated in Rodriguez et al. (2004):

> "This paper presents numerical simulations of flow through a natural meandering river using two different models: a depth-averaged numerical code with secondary flow correction and a fully 3-D, state-of-the-art Computational-Fluid-Dynamics (CFD) code. The scale of the problem examined here is in the overlapping domain of applicability of the two models, providing a basis for moving from these simulations both to larger- and smaller-scale analyses. The 2-D model generates only depth-averaged results and can be applied to large spatial domains, while the 3-D model generates a complete and detailed picture of the flow field, but at the expence [sic] of considerable computational time. In fact, the numerical simulation is rather ambitious for application of a full 3-D CFD model, as demonstrated by the lack of previous simulations at the scale

> and degree of complexity of the reach examined in this study. Neither of the models has been applied at the field-scale using data for a natural meandering stream."

It is concluded, therefore, that the approach described in Rodriguez et al. (2004) is not appropriate for application to the Housatonic River modeling study. Please refer also to the response to General Issue 3.

References:

Rodriguez, José F., Fabián A. Bombardelli, Marcelo H. García, Kelly M. Frothingham, Bruce L. Rhoads, and Jorge D. Abad. 2004. High-resolution numerical simulation of flow through a highly sinuous river reach. *Water Resources Management* 18: 177-199.

Frank Gobas:

The report documents that model calibration has been carried out to a significant degree in the hydrodynamic, sediment, chemical fate and bioaccumulation models. In my personal view, the model development has embraced calibration a little too strongly at the expense of evaluating model performance and model uncertainty. The calibration of the model has produced a model that has reasonable central tendencies and produce reasonable values for mean conditions such as flow rates, TSS and PCB concentration on TSS and in biota. However, the uncertainties in the model predictions require further attention before the model can be used productively to explore remedial options.

RESPONSE 6-FG-1:

More information on model performance will be presented during the Phase 2 Calibration and Validation. Please refer also to the response to General Issue 11 and to Response 4-FG-2 (reliance on model calibration).

In terms of additional work, it is possible to collect new PCB concentration data sets to carry out a model performance analysis that is not dependent on the collected data. Alternatively, it may be possible to revisit existing data sets and calibrate the model to certain data while using other available data for model performance analysis and uncertainty analysis. A more daring approach is not to use PCB concentration data at all in the model calibration phases. This should be possible for the PCB fate and bioaccumulation model.

RESPONSE 6-FG-2:

The data collection phase of the project has concluded. EPA believes that the recently completed human health and ecological risk assessments demonstrate elevated risks at this site, and that it is prudent to proceed with the remainder of the modeling study. Accordingly, with very limited exceptions, no additional data will be collected.

The modeling approach has been modified (see response to General Issue 2) in response to comments from the Reviewers; the most recent data will be retained for use in evaluating model performance for validation.

One area where the model calibration is lacking is in the temporal behavior of the PCB concentrations in sediment and biota in the River. This characteristic could not be calibrated very well because PCB concentrations did not show significant changes over time during the study period. As a result, there is little information on the performance of the model in terms of predicting future PCB concentrations in response to remedial options. There is not a simple solution to this problem. One approach that could be pursued is to better characterize some key loss processes of PCBs in the River. This would involve characterizing PCB degradation rates and volatilization rates. These rates may have a significant effect on the temporal response of PCB concentrations in the River. Although this work would not actually test the temporal response of the model, the credibility of the model would be improved by a better presentation of mechanisms of chemical loss.

RESPONSE 6-FG-3:

EPA agrees that the temporal gradients in PCB concentrations are weak over the calibration period, and that the sensitivity of model parameters may change during longer-term simulations; however, given the data, longer-term calibration and validation may not satisfy this concern because of a lack of clear temporal gradients in the data over the period of record. Please see also the response to General Issue 2.

Please refer to the response to General Issue 5 regarding volatilization, and Response 3-FG-1 regarding PCB degradation.

Despite the large amount of effort that has been devoted to modeling and data collection, I am not convinced that, at this point, a holistic understanding of the fate of PCBs in the River has emerged. The report is unclear about what the key processes are controlling the fate of PCBs in the River.

RESPONSE 6-FG-4:

The processes controlling the fate and transport of PCBs in the Housatonic River system have been discussed in the Modeling Framework Design (WESTON, 2004a), the RFI Report (BBL and QEA, 2003) and the Model Calibration Report (WESTON, 2004b). Although the understanding of these processes may not be perfect, EPA believes that the important fate-controlling processes have been identified and are sufficiently well understood to support achieving the goals of the modeling study.

One, in my view, useful approach is to add PCB flux diagrams to the report. Flux diagram are a useful and simple tool to integrate a lot of information with the goal to determine the controlling processes in the River.

RESPONSE 6-FG-5:

EPA agrees that summary figures are useful for integrating substantial amounts of information and communicating the important features of the system. The development of process-based flux diagrams for inclusion in the Model Validation Report will be considered. Although not broken down by processes, the PCB mass flux diagrams in the Model Calibration Report (Figures B.4-45 through B.4-47) were included to synthesize substantial amounts of information and summarize the transport of PCBs through the PSA during the calibration period.

Wilbert Lick:

The model has serious deficiencies as far as the descriptions of (a) erosion, especially at high shear stresses, (b) flocculation and deposition, (c) PCB flux and depth of mixing layer in the river, and (d) PCB flux and depth of mixing layer on the floodplain.

RESPONSE 6-WL-1:

Please refer to the responses to General Issues 6, 7, and 8, and, with regard to (d), Response 3-WL-34.

In addition, the spatial scale of the model in the river is inadequate to even approximately describe the variability of erosion/deposition in the river and hence any remediation activities in the river.

RESPONSE 6-WL-2:

Please refer to the response to General Issue 4.

Suggestions for improvements are as follows.

(a) Erosion at high shear stresses. Because of the low values of n used in the erosion equation, the erosion rate at high shear stresses and big events is not predicted properly. A re-examination of the data with a grouping of the data for similar sediments and similar depths and use of Eq. (5) should be helpful. In any event, the fact that $n \sim 2$ or more is a very strong experimental fact and hence must be considered seriously.

RESPONSE 6-WL-3:

Please refer to the response to General Issue 7.

Additional testing by means of Sedflume on well-mixed sediments would be useful. This would be relatively simple and would be as follows. Test three types of sediments (each well-mixed for uniformity within the core, one typical of coarse-grained sediments and one typical of finer-gained sediments from the main channel and one from Woods Pond). For each type of sediment and for two to three consolidation times, measurements of erosion rate as a function of shear stress should be made. This would show the effects of consolidation (which is needed in the model) and also demonstrate more clearly whether n is approximately two.

RESPONSE 6-WL-4:

The data collection phase of the project has concluded. EPA believes that the recently completed human health and ecological risk assessments demonstrate elevated risks at this site, and that it is prudent to proceed with the remainder of the modeling study. Accordingly, with very limited exceptions, no additional data will be collected.

The formulation of bed armoring needs to be improved. This can be done by introducing a thin active layer (a few particle diameters thick) which is due to physical mixing at the sediment-water interface.

RESPONSE 6-WL-5:

Please refer to Response 3-WL-11.

(b) Flocculation and deposition. As described in the response to Question 3, a simple, time-dependent model of flocculation is available and should be used. This, along with experimental results on settling speeds, should give a better and more rational description of settling as compared with making up an equation for settling.

RESPONSE 6-WL-6:

Please refer to the response to General Issue 7.

(c) PCB flux and depth of mixing layer in river. Based on existing information, the depth of the mixing layer is much less than 6 inches (15 cm) and is probably more like 2 to 4 cm. A thorough review of the literature and on-going work should be done to ascertain this. Some evidence of this layer from field observations and measurements on the Housatonic would be useful. Some evidence of number, type, and activity of benthic organism is needed. Field tests of PCB diffusive fluxes from the sediment to the overlying water should be made. These have been done elsewhere, so that it's not a new procedure.

RESPONSE 6-WL-7:

Benthic community analyses, including identification, enumeration, and life history categorization of infaunal and epifaunal organisms, were conducted for the study area as part of the ecological risk assessment. This assessment has included both a thorough evaluation of the literature and consideration of site-specific data. Please refer also to the responses to General Issues 6 and 8.

As discussed in the response to Question 3, the magnitude of the diffusive flux needs to be re-examined. Its value is not justifiable.

RESPONSE 6-WL-8:

Please refer to the response to General Issue 8.

In the modeling, the concept of a well-mixed layer should be abandoned; it is not correct physically and it is not necessary mathematically. It should be replaced by a diffusion with non-equilibrium sorption model (Lick et al., 2004) which is more realistic, more accurate, and requires little extra computational time (hardly noticeable).

RESPONSE 6-WL-9:

As is typically the case for the widely used contaminant fate models, EFDC incorporates a finite difference approximation for the vertical profile of a contaminant in sediment. Thus, the surface layer and each of the other layers is, by definition, completely mixed in a mathematical sense (i.e., the contaminant concentration within each layer is a single number). In addition, there is particle mixing between layers that are within the upper 15 cm of the bed (which was specified during calibration), and there is also diffusion of dissolved (i.e., in pore water) contaminant throughout the entire depth of the sediment bed. However, the net effect of these processes is not necessarily a well-mixed layer.

The equilibrium-based approach is routinely used to represent conditions in sediment where the particle residence time is typically relatively long in comparison to sorption and desorption time scales.

The use of a non-equilibrium approach represents an interesting alternative to the more common approach used in contaminated sediment studies such as the Fox River and Hudson River studies (Velleux and Endicott, 1994; Connolly et al., 2000, respectively). Such a major shift in the modeling approach is not possible at this point in the project. The approach implemented by EPA is consistent with the MFD and the Reviewer's comment made during the MFD Peer Review:

> "Because of these difficulties, the contaminant flux (except for resuspension/deposition) must be modeled by use of a bulk mass transfer coefficient acting over some length scale. These are both empirical parameters to be determined by calibration and should be labeled as such. There should be no pretense that somehow these processes are being modeled from basic principles." (W. Lick; May 17, 2001)

References:

Connolly, J.P., J.A. Zahakos, J. Benaman, C.K. Ziegler, J.R. Rhea, and K. Russell. 2000. A model of PCB fate in the Upper Hudson River. *Environ. Sci. Technol.* 34(19):4076-4087.

Velleux, M. and D. Endicott. 1994. Development of a mass balance model for estimating PCB export from the Lower Fox River to Green Bay. *J. Great Lakes Res.* 20(2):416-434.

(d) PCB flux and the depth of the mixing layer on the floodplain. Although some work has been done on sediment-water fluxes in a river, sediment-water fluxes on the floodplain are quite different in character. A more detailed investigation of this process is necessary including (a) field tests of the diffusive flux and (b) some evidence for a mixing layer. What is the cause of the flux and is there a mixing layer?

RESPONSE 6-WL-10:

The relative importance of these two processes on the floodplain was discussed in Responses 3-WL-34 and 3-WL-35. In addition, note that these processes are active only during times when the floodplain is inundated; therefore, they are of relatively minor importance in comparison to processes in the channel.

(e) Grid sizes and number of cells in river. Because of the dissimilar amounts of erosion in the deeper parts and erosion/deposition in the shallower parts of the cross-section of a river, a minimum of three cells across the river (two shallow, near-shore cells and one deeper, center cell) should be used in the calculations. This of course increases the computational time.

RESPONSE 6-WL-11:

Please refer to the response to General Issue 4.

However, the computational time can be greatly decreased by (a) separating the hydrodynamic and sediment transport calculations, (b) for small and moderate flows, approximate and calculate the hydrodynamics as a sequence of steady-state flows at discrete values, (c) do similar approximate calculations for sediment transport, and (d) only treat big events in detail. This will greatly decrease the computational time – more than sufficient to offset the increase in the number of grid cells. It might also be worthwhile to increase the length of the grid cells and hence decrease the number of cells along the river.

RESPONSE 6-WL-12:

Please refer to the response to General Issue 3.

It is not clear to me at this time how the model deals with erosion/deposition together with molecular diffusion and bioturbation. To do this properly, the sediment bed should be vertically divided into layers no more than a centimeter thick with additions and subtractions of mass from the surface layer as erosion/deposition occurs. Adding on the order of 10 to 20 layers of this type in the sediment bed does **not** appreciably increase computational time. These are layers which have essentially no computations associated with them and are only there until needed; in essence, their presence only requires minimal bookkeeping and no significant computation.

RESPONSE 6-WL-13:

The Reviewer's interpretation of a layered bed with additions or subtractions from the surface layer due to deposition or erosion is correct. Bioturbation is represented in EFDC by vertical mixing of contaminants between adjacent sediment layers to a maximum depth, which is specified through model input, within which benthic mixing is assumed to occur. Molecular diffusion is represented by vertical mixing of contaminant concentrations in pore water of adjacent sediment bed layers.

However, it is not correct to assume that there is a minimal effect on computational time associated with an increase in the number of vertical sediment layers. An implicit matrix solution scheme computes changes in concentration in each layer in the entire vertical column of sediment layers of a particular grid cell. The efficiency of the solution scheme would be lost if conditional ("if") statements were added to control the program-logic to segregate the solution into zones based on biological mixing. In tests of the effect of using 20 vertical layers in the sediment, the total simulation time increased by 40% because "virtual memory" was required to accommodate the additional array sizes associated with the increased vertical segmentation in the sediment. The associated disk input/output (I/O) caused the substantial increase in runtime.

These suggestions are not radical, untried, unproved, new ideas. A sediment and contaminant transport model, SEDZLJ, already exists (Jones and Lick, 2000, 2001, 2002; Lick et al., 2004) which incorporates most of these ideas. It uses Sedflume data and includes multiple size classes, a unified treatment of suspended load and bedload, bed armoring, an active layer due to physical mixing at the surface, HOC flux due to molecular diffusion and bio-diffusion as well as transport, and fine layering in the vertical in the sediment bed to adequately describe sediment bulk properties and HOC flux. SEDZLJ is presently being incorporated into EFDC.

RESPONSE 6-WL-14:

Although it appears to be the case that SEDZLJ is well developed, not all of the ideas discussed above (i.e., only "most of these ideas") are included in the present version. Thus, further development and testing would be required before the preceding features would be fully operational, and it is uncertain that the implementation of this alternative approach would be successful without further refinements to the model.

As recently as February 2005, SEDZLJ was still a one-dimensional model and not available for distribution (Craig Jones, personal communication, 2005). The author of SEDZLJ has been working to incorporate SEDZLJ into an earlier version of EFDC, however, as of August 2005 this code was still being tested (Jesse Roberts, personal communication, 2005). At this point, EPA does not believe that it is practical to extract the SEDZLJ code from an alternate version of EFDC and incorporate it into the version of EFDC being applied to the Housatonic River. Substantial code changes would be required and would need to be tested, introducing unacceptable delays in the project.

Any state-of-the-art model (such as EFDC) that is applied to a complex system such as the Housatonic River must undergo continual testing and development. It will always be possible, in hindsight, to identify more recently developed models or formulations that may provide new and useful features, but have not undergone the rigorous testing that would be required for implementation within the modeling study.

Further discussion related to this general subject matter is included in the response to General Issue 3.

References:

Jones, C. 2005. Personal communication (e-mail) from Craig Jones, Sea Engineering, to Edward Garland, HydroQual, re: response to inquiry about availability of computer model, SEDZLJ.

Roberts, J. 2005. Personal communication (e-mail) from Jesse Roberts, Sandia National Laboratories, to Edward Garland, HydroQual, re: status of efforts to incorporate SEDZLJ in EFDC.

E. John List:

Overall I am impressed at the depth and breadth of the work that has been completed and feel much more positive about the outcome than I did four years ago. The shortcomings in the modeling that have been identified here, and by others on the Peer Review Committee, can most probably be rather easily overcome. There is some additional calibration work that is necessary before the model can be used to predict future conditions with any degree of confidence. The issues that I see that need resolving are:

1. Attempt to resolve why PCB concentrations within the river reach have such extreme spatial variability on a horizontal scale. Is it because the PCB was initially released as free product and formed droplets that were carried by the stream and deposited out of the water column, or is it some other reason? Does the spatial inhomogeneity of PCB in the sediments reflect spatial inhomogeneity of organic carbon, and if so, why is the organic carbon distribution spatially inhomogeneous? Now that no further new releases of PCB are occurring will the erosion-deposition processes in the river lead to homogenizing the PCB concentrations in the sediment? If not, what is the mechanism by which the inhomogeneity is being maintained? How can this mechanism be included in the fate and transport model? If it cannot be readily incorporated is it still possible to use such a model to describe the fate and transport of the PCB? It is not a satisfactory response to simply ignore the issue. Given that the modeling exercise is chartered with predicting the outcome of remediation strategies it is my opinion that the fate and transport modeling process applied so far, which focuses only on spatially-averaged concentrations in a situation where there is extreme spatial variability, is inappropriate. The failure of the study team to investigate and understand the basis for this extreme, and apparently temporally sustained, spatial variability in concentration, is somewhat disturbing, especially when it appears to be of such importance to understanding the fate and transport of the PCB.

It is not sufficient to simply acknowledge this deficiency with the statement (pg. 5-58): “This variability, which is a combination of stochasticity and analytical variability, is not represented in the model inputs, and therefore it is not represented in the model output.”

RESPONSE 6-JL-1:

Although there is small-scale spatial variability in sediment TOC, it is not sufficient to explain the more pronounced variability in PCB concentrations, i.e., carbon-normalized PCB concentrations are not uniform throughout the PSA. As discussed in the response to General Issue 5, although it may be desirable to understand such small-scale processes, EPA does not believe such an understanding is necessary to achieve the goals of the modeling study.

EPA's approach is consistent with observations from the same Reviewer during the Peer Review of the Modeling Framework Design:

> "The model that is proposed will attempt to describe the rates of erosion and deposition on a 20-meter grid plan. . . . It is implausible to think that the riverbed sediments can be characterized on a grid scale this small, so that attempting to model the fate of the sediment on such a scale appears quite inappropriate. In any case, the model output is to be aggregated . . . on a grid scale that is about 250 times as large." (J. List, May 22, 2001)

Please refer also to the response to General Issue 1.

2. Representation of a section of river by a single element that averages over the lateral extent of the basic channel is unlikely to provide the resolution of erosion and deposition that is required. In my opinion at least three basic elements are needed to represent the potential for erosion and deposition to occur at the same river mile, which is what really happens in most streams with which I am familiar.

RESPONSE 6-JL-2:

In his observations during the MFD Peer Review, the same Reviewer indicated that a single element across the lateral extent of the channel would be adequate:

> "It therefore would appear to make much more sense to use these data to calibrate a transport model that is based upon a one-dimensional representation of the river system." (J. List, May 22, 2001)

> and

> "The basic problem is not with the theoretical rigor of the equations, but with the context within which they are placed. For example, the description of resuspension and erosion of particles can be described quite adequately by using the empirical data developed by the SEDFLUME apparatus. The issue becomes how to use these data in the modeling when it is known from the sediment sampling in the river channel and flood plains that the sediments are extremely variable with respect to the rate of erosion. It is not possible to describe completely the surface and depth distribution of the sediment properties that control erosion at the fine

> scale necessary to apply a two-dimensional model with a 20-meter (or less) grid scale." (J. List, May 22, 2001)

Please refer also to the response to General Issue 4.

3. The issue of why the modeling does not adequately reproduce the variance of tPCB concentration in the water column and in the biota needs resolving. If the model cannot resolve the variance of the distributions of PCB it is difficult to see how any statistical hypothesis regarding the relative efficacy of different remediation strategies can be properly resolved by the output of the modeling.

RESPONSE 6-JL-3:

EPA does not agree that it is necessary to resolve the variance of the distributions of PCB to achieve the goals of the modeling study. As the Reviewer observed during the MFD Peer Review:

> ". . . from the sediment flux data that have been developed in the field it should be possible to give average sediment properties that can be used to describe in a general way the resuspension of river bed sediments and flood plain sediments. This is not unusual in fluid mechanics; sometimes less is more." (J. List, May 22, 2001)

Please refer also to the responses to General Issues 1 and 10.

4. Some basic statistical analysis on both the data input to the EFDC model and its output appears not to have been completed. In the absence of this analysis it is not clear if variations in the input data (e.g., mean PCB concentrations in the sediment elements along the river) have any statistically significant variation from one end of the 11-mile river reach to the other.

RESPONSE 6-JL-4:

Please refer to Response 4-JL-3.

The predicted fluxes of PCB and sediment similarly appear to have no uncertainty analysis. These are serious omissions that can easily be corrected.

RESPONSE 6-JL-5:

Please refer to the response to General Issue 11.

5. The apparent bias in the food chain model in Reach 5D needs resolution. In the absence of any field data in this reach it is difficult to make an unequivocal judgment that it is actually model bias, but the results are so far out of congruence with the other reaches that it is very suggestive of something amiss in the modeling.

RESPONSE 6-JL-6:

Please refer to Response 2-JL-1.

ADDITIONAL COMMENTS

Douglas Endicott:

I think a good technical writer could do a lot with this report in terms of improving readability and clarity.

RESPONSE AC-DE-1:

EPA recognizes that the Model Calibration Report (particularly the technical appendices) presents a large amount of highly technical information. In fact, the document did undergo careful and iterative editing to improve consistency, clarity, and readability.

EPA will direct more effort toward streamlining and reformatting the Model Validation Report to improve readability.

PCB Fate and Transport Schematic:

I usually take the time to examine model schematics, and I find them particularly valuable in understanding how the conceptual model is applied. In this case (Figure B.4-1) there are a number of errors in the schematic which should be corrected. These errors should be obvious to the modeling team.

RESPONSE AC-DE-2:

A revised version of the schematic is presented as Figure AC-DE-2. The bioaccumulation portion of the original figure has been eliminated because those calculations are performed in FCM.

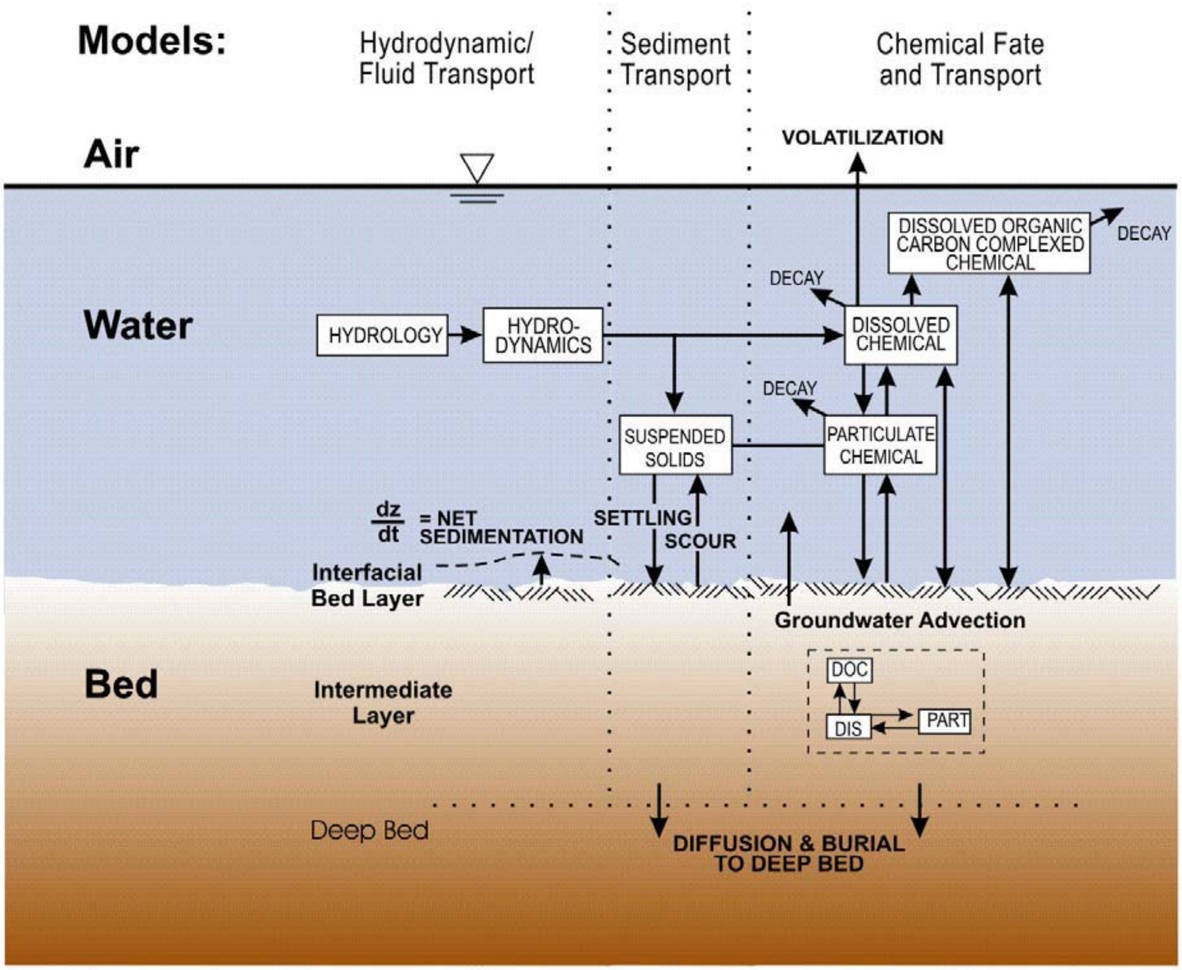


Figure AC-DE-2 Schematic of EFDC and Linkage of Hydrodynamic/Fluid Transport, Sediment Transport, and Fate and Transport Submodels

Bioaccumulation of coplanar PCBs:

As previously commented, I don't believe that the "correction factor" approach used in the food chain model to reduce the bioavailability of coplanar PCBs is scientifically correct. Based primarily on comparisons of the highest-quality measured PCB BSAFs for fish to model predictions assuming no metabolism (Burkhard et al., ES&T 2004), it appears that congener 77 and possibly congener 126 to a lesser extent are very slightly reduced through metabolism. Since the congener-specific metabolism would probably be aryl hydrocarbon receptor (AHR) mediated, metabolism in invertebrates is unlikely (i.e., it should only be taking place in the FISH). One cannot rule out a bioavailability effect associated with affinity of more planar PCBs with small amounts of black carbon in water and sediments. I think this should be addressed in greater depth in the Bioaccumulation section of the Calibration report.

RESPONSE AC-DE-3:

EPA believes that there are multiple potential explanations and uncertainties associated with the observed phenomenon of lower bioavailability of certain coplanar PCB congeners. Although metabolism was identified as a potentially relevant process in the Model Calibration Report, EPA did not intend to imply that metabolic clearance was the only explanation for the observed differences. The approach used in the Model Calibration Report does not include an assumption about the specific mechanism by which reduced uptake occurs. As such, the derivation of an empirical scaling factor based on fish BSAFs (i.e., comparisons of fish:sediment concentration ratios for coplanar congeners versus ratios for recalcitrant congeners) does not discriminate between differences in bioavailability for prey (including invertebrates) or fish.

Because no assumptions regarding the mechanism resulting in the reduction in bioaccumulation for certain congeners were made, the application of the empirical scaling factors to the base of the food web is a reasonable approach that results in acceptable model calibration for these congeners in fish. The uncertainty in invertebrate PCB concentrations is greater, however, and the Reviewer is correct that these uncertainties should have been discussed in greater detail in the Model Calibration Report.

EPA agrees that if congener-specific differences are mediated only by the Ah receptor, the simulations of invertebrate tissue concentrations would be biased low by FCM for the congeners in question. On the other hand, if congener-specific differences are mediated by the bioavailability effect associated with the affinity of more planar PCBs to small amounts of black carbon, reduced PCB bioaccumulation in invertebrates would be expected. Based on examination of simulated versus measured coplanar concentrations in benthic invertebrates (Figure C.3-16), it is unclear whether the model should adjust coplanar bioavailability at the base of the food web (affecting all organisms), or only in fish and not in invertebrates.

EPA acknowledges that incorporating congener-specific correction factors for coplanar congeners introduces some additional uncertainty. EPA does not believe there is sufficient information in the literature to mechanistically specify processes of reduced uptake related to metabolism, black carbon, or other congener-specific factors with any more certainty. As such, the invertebrate simulations are more uncertain for these congeners than the simulations for fish, and should be either excluded from the report, or the uncertainties discussed in more detail in the report.

Even with these uncertainties, EPA believes that the evaluation of model performance for specific congeners in fish (with or without bioavailability adjustments for coplanar congeners) provided a useful test of model robustness for this class of contaminants.

PCB elimination by fish:

Biphasic (multicompartment or deep storage) elimination is sort of a hidden feature within the QEA foodchain model. By that, I mean the literature documenting the development and application of this model does not address this fairly important feature (as far as I am aware).

RESPONSE AC-DE-4:

Section 6.2.5.4 of the Model Calibration Report discusses biphasic elimination under the heading “Contaminant Elimination in Fish.” This brief summary also points the reader to Section 1.4.4 of the “Detailed Model Description” (Attachment C.1, pages 15-18). These four pages describe the rationale for the biphasic elimination model in more detail. Several lines of evidence to support the model are explored and numerous references are provided. This material was obtained from the QEA Hudson River model report (QEA, 1999).

Reference:

QEA (Quantitative Environmental Analysis, LLC). 1999. PCBs in the Upper Hudson River. Volume 2 - A Model of PCB Fate, Transport, and Bioaccumulation. Report prepared for the General Electric Company, Albany, NY. Amended July 1999. URL: http://www.ge.com/files/usa/en/commitment/ehs/hudson/Volume_2.pdf

Again, this should be addressed in greater depth in the Bioaccumulation section of the Calibration report. In addition, computed elimination rates should be compared to rates measured in fish.

RESPONSE AC-DE-5:

As part of model calibration, EPA did conduct tests of rate parameters (e.g., elimination rate tests) and gauged the model output against values obtained from the literature. EPA will consider documenting some of these tests in the Model Validation Report. Please refer also to Response AC-DE-4 above (documentation of biphasic elimination).

ATTACHMENT 1

FRESHWATER BIOTURBATION DEPTH

TABLE OF CONTENTS

LIST OF TABLES

Table No. **Page**

LIST OF FIGURES

Figure No. **Page**

ATTACHMENT 1

FRESHWATER BIOTURBATION DEPTH

1. INTRODUCTION

1.1 OBJECTIVES

The EFDC model (and the linkage between EFDC and FCM) requires specification of the depth of the surface sediment layer(s) that exchange PCBs with overlying water via the processes of advection and/or diffusion (variously termed the "active," "well-mixed," "bioavailable," or "bioturbation" layer).

The objectives of this review were to:

- Estimate the sediment depths over which biological mixing of sediment occurs in Housatonic River sediment (i.e., bioturbation depths).
- Estimate how the bioturbation depths vary as a function of substrate/habitat type within the Primary Study Area (PSA).
- Determine the depth of the surface sediment layer that best estimates the average exposures of benthic organisms in the food web. This depth is related to the bioturbation depths described above.

No direct site-specific measurements of bioturbation depths have been made in the PSA. Therefore, estimates of bioturbation depths were based on literature studies that were screened for relevance to the Housatonic River. Characterizations of physical substrate and biology of the Housatonic River (e.g., Ecological Characterization presented in the Ecological Risk Assessment [WESTON, 2004a]), sediment profiling, and invertebrate taxonomy and life history were considered in the review.

1.2 TERMINOLOGY

The evaluation of bioturbation requires definition of several terms. The definitions below are intended to specify how these terms are used in this document:

- *Bioturbation* – The displacement and mixing of bed sediment particles by benthic organisms. All animals that feed, burrow, move, or excrete in or on the sediment cause bioturbation (Campbell et al., 1988). Bioturbation is defined to include biologically mediated sediment mixing, and excludes abiotic processes such as sediment erosion and deposition.

- *Bioturbation depth* – The total depth over which bioturbation occurs. This total depth has been conceptualized as two sublayers, including the *biologically mixed layer* (i.e., shallow thorough mixing) and the *biologically influenced layer* (i.e., deeper, less intense, mixing) (see Figure 1).

- *Sediment mixing* – The rearrangement of sediment particles from all processes, including both bioturbation and physical fate and transport processes, such as sediment erosion and deposition.

- *Well-mixed layer* – The layer of sediment that lacks a consistent vertical profile in PCB contamination due to sediment mixing.

- *Well-mixed depth* – The depth of the well-mixed layer within the sediment bed.

- *Biologically mixed layer* – The surface sediment layer in which bioturbation results in substantial reworking of sediment particles. The depth of this layer is significantly influenced by the maximum depth to which benthic organisms feed, because feeding activities generally play the greatest role in sediment reworking (see Figure 1).

- *Biologically mixed depth* – The depth of the biologically mixed layer within the sediment bed.

- *Biologically influenced layer* – The subsurface sediment layer over which bioturbation occurs, but where infaunal densities are lower and rates of sediment reworking decrease. The depth of this layer is generally greater than that of the biologically mixed layer described above, because benthic organisms occur at depths greater than their maximum feeding depth, and exert some physical influence on sediment at these depths. Therefore, although the greatest magnitude of sediment mixing is observed within the zone of active feeding, bioturbation can also occur from other processes (e.g., biodiffusion).

- *Biologically influenced depth* – The depth of the biologically influenced layer within the sediment bed.

- *Bioavailable depth* – The depth of the surface sediment layer that best approximates the average sediment PCB exposures of infaunal invertebrates. Depending on the vertical distribution of organism densities, bioavailable depth can differ from the bioturbation depth.

2. LITERATURE REVIEW

2.1 METHODS

To identify relevant information on infauna-mediated vertical mixing in sediment, a detailed literature review was conducted. Extensive bibliographies relating to bioturbation and sediment flux were compiled by EPA contractors, and supplemented with information provided by General Electric (GE) (QEA, 2005). These bibliographies, which included a total of 139 documents, were the starting point for the literature review.

For each of the 139 documents, the following evaluation procedure was applied:

Document Screening – Documents that did not meet the minimum requirements for relevance were screened out. Documents eliminated from the review included cases for which:

- The document lacked information regarding depth of sediment (e.g., studies of flux of contaminants from sediment, studies limited to lateral mixing).
- The document described the development of a mathematical model for bioturbation but without accompanying field or laboratory data.
- The document described a benthic study unrelated to sediment mixing or the influence of biota on sediment (e.g., studies on bioaccumulation, life cycles, bioenergetics, etc.).

Data Summary – In tabular format, relevant data were extracted, including:

- Sediment substrate characteristics (e.g., grain size, TOC).
- Types of organisms investigated.
- Physical environment (e.g., marine vs. freshwater, in situ study vs. microcosm).
- Mixing depth (either directly reported or inferred).
- Other relevant data (e.g., vertical distribution of benthos, depth-related influence of biota on sediment).

Relevance Ranking – For each study retained following the screening stage, a qualitative assessment of relevance to the Housatonic River PSA was assigned. The quality and relevance of each document was categorized as "high," "moderate," or "low" based on the following criteria:

- High Relevance – A document was categorized as having high relevance if it considered a freshwater environment and met all of the following criteria: (1) contained data on vertical mixing; (2) sediment characteristics (e.g., grain size, TOC) were provided and approximated those in the PSA; (3) the types of organisms responsible for the sediment mixing described in the paper were similar to those in the PSA. The requirement for freshwater studies was based on the observation that biologically mixed layers in marine systems tend to be slightly deeper than in rivers and lakes (Clarke et al., 2001).

- Moderate Relevance – A document was categorized as having moderate relevance if it considered a freshwater environment and did not satisfy all the criteria for high relevance, but satisfied one of the following conditions: (1) included data on vertical mixing but insufficient information to determine whether the sediment type resembled that of the PSA; (2) vertical mixing depth in sediment was not directly reported, but the information provided was sufficient to infer the biologically active depth. Examples of the latter included the following: studies reporting the maximum depth at which a chemical tracer compound was released to the water column following sediment mixing; information on the physical influence of benthos on pore water redistribution; and information on maximum feeding depths of benthos.

- Low Relevance – Documents that did not meet the criteria for classification as moderate or high relevance, but that met one or more of the following criteria were categorized as having low relevance: (1) the document provided information on the vertical distribution of benthic organisms but did not provide direct information on the influence of the organisms on sediment mixing; (2) the document described bioturbation in an environment of limited relevance to the Housatonic River (e.g., very coarse sediment, marine or estuarine environments). Marine data were included in the literature review (but categorized as having low relevance) because there is some correspondence between marine and freshwater benthic communities in terms of distribution and intensities of bioturbation processes (Clarke et al., 2001; McCall and Tevesz, 1982; Rhoads and Boyer, 1982).

Identification of biologically mixed depths was made using a combination of directly reported results from the documents and inferences (using professional judgment). Greater use of judgment was required for interpretation of studies with low to moderate relevance.

2.2 RESULTS

2.2.1 General Findings

Several general themes emerged from the literature review that were useful for evaluating specific studies and for inferring biologically mixed depths when they were not reported directly.

The depth to which bioturbation occurs is a function of the burrowing depth of resident organisms, the density of the organisms in the substrate, and the behaviors of diverse assemblages of benthic organisms and their interactions with the physical environment (Schaffner et al., 1997; Wheatcroft et al., 1994). Bioturbation is primarily the result of feeding and burrowing behaviors of infaunal and, to a lesser degree, epifaunal benthos. Bioturbation by most benthos is generally characterized by random mixing. However, mixing by tubificid oligochaetes is directional due to a feeding behavior characterized by ingestion of sediment at depth within the sediment bed and egestion at the sediment-water interface (McCall and Fisher, 1979).

A variety of macrobenthos play a role in bioturbation, although their contributions may vary due to differences in factors such as feeding behaviors and burrowing depth. Tubificid oligochaetes are often considered some of the most effective bioturbators due to their feeding behavior and maximum feeding depths (generally on the order of 10 cm). Freshwater insects (including chironomids) and amphipods are also considered important bioturbators, although they tend to influence a smaller depth range within the sediment bed (Clarke et al., 2001). Bivalves may also play a significant role in bioturbation with some groups (e.g., Unionidae; large freshwater clams) influencing more sediment than others (e.g., Sphaeridae; fingernail or pea clams).

Bioturbation depth increases with increasing biomass density due to increased competition for food sources and associated movement. Although some invertebrates burrow deeper than 10 cm, bioturbation rates are generally considered to decrease with increasing depth in the sediment bed because organism densities generally decline with increasing depth (Bromley, 1996).

Clarke et al. (2001) conceptualized vertical profiles of biological activity as having multiple distinct bioturbation zones (Figure 1). A surficial zone of intensively reworked sediment is biologically mixed over relatively short time scales. This layer, which is continually mixed, is defined herein as the "biologically mixed layer." Beneath the biologically mixed layer, a mid-depth zone is characterized by decreasing biological activity with increasing depth. This layer is defined herein as the "biologically influenced layer." Deeper zones are poorly understood in comparison with the overlying zones and are assumed to have negligible to low concentrations of biological mixing. Clarke et al. (2001) recommended that, at a minimum, bioturbation should be treated as a two-layer system with an overlying continually mixed layer and an underlying bio-diffusion layer.

The depth of the upper biologically mixed layer has been linked to the activity levels of invertebrates, and in particular, the depth at which active feeding occurs. Sediment mixing by oligochaetes is predominantly a function of their feeding behavior rather than mixing caused by movement (burrowing). Therefore, in sediment where oligochaetes are abundant, the depth of the biologically mixed layer often corresponds to the maximum feeding depth, which is generally less than the maximum depth to which benthos frequently burrow (Cunningham et al., 1999; Fisher et al., 1980).

2.2.2 Specific Findings

Of the 139 papers reviewed, 43 passed the preliminary screening stage. Of these, 7 were ranked as having high relevance, 13 were ranked as having moderate relevance, and 23 were ranked as having low relevance. The papers that failed the preliminary screening are not included in Table 1 or the list of references. Table 1 summarizes the relevant data from each document. As expected, the individual documents reported a wide range of biological mixing depths, depending on the physical and biological environments evaluated.

Several of the bioturbation documents consisted of reviews of multiple freshwater environments and offered general guidance for selection of biological mixing depths, including:

- Campbell et al. (1988) emphasize the importance of freshwater oligochaetes for bioturbation, and indicate that bioturbation effects are "most pronounced" in the upper 6 cm (2 inches) of sediment, with little or no bioturbation below 20 cm (8 inches).

- Clarke et al. (2001) identified an upper layer subject to frequent and thorough mixing by shallow bioturbating organisms. This surficial zone, associated with the transition between oxic and anoxic sediment, may extend to a depth of 10 cm (4 inches) or more.

- Fisher (1982) indicated that bioturbation depths for amphipods and sphaeriid clams are limited to the top few centimeters, but oligochaetes and unionid clams mix sediment to depths of approximately 10 cm in freshwater environments.

- Krezoski et al. (1978) and Ford (1962) suggested that freshwater benthic macroinvertebrates are often found as deep as 10 cm, but are concentrated within the top 4 cm (1.5 inches) of sediment.

- EPA (2002) conducted a survey of aquatic biologists from several research facilities around the Great Lakes to evaluate bioturbation depth in a theoretical sandy freshwater sediment cap colonized mainly by chironomids and oligochaetes. Although some organisms indigenous to the Great Lakes can burrow 10 to 40 cm in soft silt or clay sediment, most of the researchers surveyed believed that most bioturbation in a sand cap would be limited to the top 5 to 10 cm.

Overall, the literature summaries suggest that bioturbation is pronounced within the top 5 to 10 cm of sediment at most sites, resulting in frequent and thorough mixing of sediment in this upper layer. These generic bioturbation estimates were combined with more specific studies (summarized in Table 1) and site-specific assessment of the Housatonic River (Section 3) to derive site-specific estimates of bioturbation depths (Section 4).

3. SITE-SPECIFIC CONSIDERATIONS

3.1 HABITAT/SUBSTRATE CHARACTERIZATION

The PSA includes sediment with different physical and biological characteristics that mediate bioturbation potential. The main channel PSA benthic communities exhibit distinct differences between upstream and downstream habitats, as described in WESTON (2004a, b). The upstream habitats (Reach 5A) are dominated by coarser sediment particle sizes (medium sands) relative to the downstream habitats (Reaches 5B, 5C, and 6) that are dominated by silt. The backwaters (Reach 5D) represent a third habitat type, characterized by fine sediment but a different hydrological regime. The downstream reaches have higher sediment total organic carbon (TOC) content, benthic abundance/biomass, and density of aquatic macrophytes relative to upstream areas (WESTON, 2004a, b). The biological communities in the PSA reaches reflect these habitat differences, as discussed in Section 3.2 below.

Mixing depths are expected to vary within the study area based on the differences between the three general habitats. Therefore, rather than identifying a single mixing depth for the entire study area, mixing has been described for each of the three habitats. A comparison of three broad PSA environments is provided in Table 2. The portion of the PSA for which each document was considered relevant (i.e., Area 1: coarse-grained PSA sediment; Area 2: fine-grained PSA sediment excluding backwaters; Area 3: Reach 5D backwaters) is noted in Table 1.

Combining the vertical stratification of the sediment bed into biologically mixed sediment and biologically influenced sediment for each of the three habitat types results in the specification of six discrete biological mixing layers (i.e., 3 habitat types times 2 layers) for the PSA.

3.2 SITE-SPECIFIC BIOLOGICAL COMMUNITIES

Clarke et al. (2001) stressed that available knowledge on local bioturbators should supplement generic assumptions concerning bioturbation processes. Ecological characterization and taxonomic enumeration of the Housatonic River benthic invertebrate communities are of value for refining the generic bioturbation estimates cited in Section 2.2.

There are two main sources of information for the identification and classification of Housatonic River bioturbators:

- EPA conducted macroinvertebrate community structure sampling in 1999 (WESTON, 2004a, b). At each of 13 locations (12 in the Housatonic River and 1 in Threemile Pond), EPA collected 12 replicate benthic community samples using a Petite Ponar sampler. Five of these locations were located in the relatively coarse-grained substrate upstream of the Pittsfield wastewater treatment plant (WWTP). Four were located in the finer (silty) substrate between the WWTP and Woods Pond, and the remaining four stations were established outside the PSA. The sediment samples were representative of the broad substrate types found throughout the PSA.

- On behalf of GE, Chadwick (1994) investigated aquatic ecology at 10 stations in the Housatonic River, of which 3 were located in the PSA. Station HR1 was located at a "shallow water" site in the northern half of the PSA, near the confluence with Sykes Brook. Stations HR2 and WP1 were located in the southern portion (low gradient, depositional) of the PSA at "deep water" locations. A modified Hess sampler (0.1 square meters [m^2], 500 micron [μm] mesh) was used to sample erosional habitat at shallow-water stations. An Ekman grab sampler (0.02 m^2) was used at the deep-water locations.

The communities from each of these programs are summarized in detail in Attachment C.11 of the Model Calibration Report (WESTON, 2004b) and in Appendix D of the Ecological Risk Assessment (WESTON, 2004a). Information on habitat, behavior, and trophic relationships was obtained from summaries provided in Merritt and Cummins (1996), and was used to determine whether the observed organisms exhibited burrowing behaviors or were mainly associated with the sediment-water interface (e.g., clingers, climbers, and net-spinners). This information, together with measurements of total abundance and biomass, and the biomass of known bioturbators (e.g., tubificid oligochaetes), was used to assess bioturbation potential in each of the three habitat types. The following sections summarize the main findings used in the refinement of bioturbation depth estimates.

3.2.1 Coarse-Grained PSA Sediment

The Chadwick (1994) study indicated that the majority of organisms in coarse-grained sediment were dipteran larvae. Of the total composite density (4,471 individuals/square meter), 37% were sediment burrowers and 60% were surface feeders (3% could not be assigned to either category). The EPA Ponar samples also indicated that dipterans (e.g., chironomids) were numerically dominant. However, oligochaetes and bivalves were also common, comprising approximately half of the invertebrate biomass at most Reach 5A stations (WESTON, 2004a). The most abundant invertebrate (*Polypedilum calaenum*) was a grazing chironomid larva that dwells upon vascular hydrophytes, and therefore, has low bioturbation potential. However, numerous other chironomid species with high Reach 5A abundance (e.g., *Cryptochironomus* spp., *Saetheria* spp.) burrow in the sediment bed.

The overall biomass of invertebrates in the EPA Reach 5A stations was lower (by an order of magnitude) relative to downstream fine-grained habitats. Part of this difference was attributed to the higher mean PCB concentrations measured in Reach 5A; however, differences in biomass were also related to habitat differences (particularly TOC and particle size). The lower organic carbon content of the coarse-grained sediment results in a poorer source of food energy for deposit feeding organisms. As a result, the bioturbation potential in Reach 5A is expected to be lower than for downstream reaches.

3.2.2 Fine-Grained PSA Sediment

The invertebrate community at station HR2 (Chadwick, 1994) was dominated by sediment-feeding Diptera larvae, including *Dicrotendipes* sp., a burrowing chironomid that was approximately 25% of the total abundance. Tubificid oligochaete worms were also abundant, at nearly 15% of the total abundance. Overall, most of the invertebrate fauna were burrowers and can be assumed to have strong bioturbation potential. The invertebrate community in Woods Pond (WP1) also exhibited a number of potential bioturbators. The most abundant organism was the burrowing chironomid larva *Ladopelma* sp., which was approximately 40% of the total organism abundance. Other common organisms included oligochaetes and numerous burrowing/sprawling/predatory Diptera larvae.

The EPA Ponar sampling (WESTON, 2004a, b) indicated a wide range of life histories and feeding strategies. The most abundant invertebrates in the fine-grained samples were sphaeriid clams, chironomids, gastropods, and tubificid worms. Overall, the Ponar grabs indicated a significant number of burrowing organisms, including a high density of oligochaetes. These organisms are expected to have high bioturbation potential.

Overall, the bioturbation in the fine-grained habitat of the PSA is expected to be more pronounced than in the coarse-grained habitat. This is due both to the increased abundance and biomass of organisms and to the increased proportion of organisms that burrow into sediment. Furthermore, megafauna, primarily large invertebrates and fishes (e.g., carp), may be locally significant bioturbators (Atkinson and Taylor, 1991) and increase the overall rate and depth of biological mixing. Carp and goldfish, both of which burrow into sediment, comprise a significant component of the fish community in downstream fine-grained sediment, but are rare in Reach 5A.

4. APPLICATION TO THE MODEL FRAMEWORK

4.1 DERIVATION OF BIOTURBATION DEPTHS

For each of the three major habitats in the PSA, two zones of bioturbation influence were defined, the biologically mixed layer and the biologically influenced layer. Below the biologically influenced layer, bioturbation is minimal and can effectively be assumed to be zero.

Depths of biological mixing were derived primarily considering information extracted from the documents and are categorized as moderate or high relevance. The information from documents rated as low relevance was used only as corroborating evidence. For each habitat type, the subsets of documents considered directly relevant were considered as a group and professional judgment was used to estimate appropriate bioturbation depths (Table 3).

As shown in Table 3, the biologically mixed depth ranges from 4 cm in Reach 5A to 10 cm in Reach 5D backwaters. These differences among habitats are warranted based on the differences in organism abundance, biomass, and life history. On average, the biologically mixed depths in the PSA are lower than default values (i.e., 10 cm) commonly used to estimate well-mixed depths in marine systems (Boudreau, 1998; Thibodeaux and Bierman, 2003), reflecting the differences between freshwater and marine environments. There is significant uncertainty associated with the assignment of the depth and mixing rates, however, based on the variability measured within and between organisms reported in the literature.

4.2 MODEL LINKAGES

4.2.1 Derivation of Bioavailable Depth

The bioaccumulation model (FCM) requires specification of a layer of sediment that represents average sediment PCB exposures to infaunal invertebrates. As shown in Figure 1, sediment infaunal densities decrease significantly below the biologically mixed layer. PCB concentrations in the biologically mixed layer are likely to better represent average exposure conditions to invertebrates that form the base of the food web. Therefore, the bioavailable depth used for the EFDC to FCM linkage will be defined as roughly equivalent to the biologically mixed depths shown in Table 3. When EFDC-modeled sediment layers do not correspond directly to the depth intervals shown in Table 3, weighted averaging of the applicable surface sediment concentrations may be used as estimates of PCB exposure concentrations.

4.2.2 Computation of Overall Sediment Mixing

The mixing of sediment in EFDC includes both biological mixing and physical mixing due to hydrodynamic processes. The bioturbation depths specified in Table 3 will be used to specify layers of sediment in EFDC, with separate biological mixing coefficients specified for the biologically mixed layer and the biologically influenced layer, respectively. Professional judgment will be required to select mixing coefficients that simulate "thorough" and "partial" mixing in these two zones. The erosion and deposition kinetics in EFDC will introduce additional mixing that is handled separately from bioturbation.

5. SUMMARY

Based on an extensive literature review of freshwater bioturbation, and considering site-specific data on habitat and resident benthic communities, bioturbation depths were estimated for each of three PSA habitat types. The biologically mixed depth, which represents the sediment layer that is thoroughly mixed due to bioturbation, was determined to range from 4 cm to 10 cm across the PSA, depending on habitat, with deeper mixing found in depositional areas with high sediment TOC. Deeper, but less pronounced, biological mixing (i.e., biologically influenced depth) is estimated to occur at depths up to 20 cm below the sediment-water interface.

6. REFERENCES

Atkinson, R.J.A., and A.C. Taylor. 1991. Burrows and burrowing behaviour of fish. *Symposium of the Zoological Society of London* 63:133-155.

Bishop, J.E. 1973. Observations on the vertical distribution of the benthos in a Malaysian stream. *Freshwater Biology* 3:147-156.

Boudreau, B.P. 1994. Is burial velocity a master parameter for bioturbation? *Geochimica et Cosmochimica Acta* 58:1243-1249.

Boudreau, B.P. 1998. Mean mixed depth of sediments: the wherefore and the why. *Limnology and Oceanography* 43:524-526.

Brannon, J.M., R.E. Hoeppel, T.C. Sturgis, I. Smith, and D. Gunnison. 1985. *Effectiveness of Capping in Isolating Contaminated Dredged Material from Biota and the Overlying Water*. Technical Report, D-85-10. U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.

Bromley, R.G. 1996. *Trace Fossils: Biology and Taphonomy* (2nd edition), Chapman and Hall, London. p. 306.

Campbell, P.G.C., A.G. Lewis, P.M. Chapman, A.A. Crowder, W.K. Fletcher, B. Imber, S.N. Luoma, P.M. Stokes, and M. Winfrey. 1988. *Biologically Available Metals in Sediments*. National Research Council of Canada, Associate Committee on Scientific Criteria for Environmental Quality, and Committee on Marine Analytical Chemistry of the NRCC Marine Analytical Standards Program. NRCC No. 27694.

Chadwick (Chadwick & Associates). 1994. *Aquatic Ecology Assessment of the Housatonic River, Massachusetts, 1993*. Prepared for General Electric Company, Littleton, CO. May 1994.

Clarke, D.G., M.R. Palermo, and T.C. Sturgis. 2001. Subaqueous cap design: Selection of bioturbation profiles, depths, and rates. DOER Technical Notes Collection (ERDC TN-DOER-C21). U.S. Army Engineer Research and Development Center, Vicksburg, MS. http://el.erdc.usace.army.mil/dots/doer/pdf/doerc21.pdf

Cunningham, P.B., D.D. Reible, J.F. Fleeger, K.T. Valsaraj, and L.J. Thibodeaux. 1999. *Assessment of the Effects of Bioturbation in Contaminated Sediments*. Proceedings of the 1999 Conference on Hazardous Waste Research, Gateways to Environmental Solutions. pp. 276-285.

Danielopol, D.L. 1976. The distribution of the fauna in the interstitial habitats of riverine sediments of the Danube and the Piesting (Austria). *International Journal of Speleology* 8:23-51.

Davis, R.B. 1974a. Stratigraphic effects of tubificids in profundal lake sediments. *Limnology and Oceanography* 19:466-488.

Davis, R.B. 1974b. Tubificids alter profiles of redox potential and pH in profundal lake sediment. *Limnology and Oceanography* 19:342-346.

Davis, W.R. 1983. Sediment-copper reservoir formation by the burrowing polychaete *Nephtys incisa*. *Offprints from Wastes in the Ocean,* Volume II, *Dredged Material Disposal in the Ocean*. John Wiley & Sons, Inc.

EPA (U.S. Environmental Protection Agency). 1987. *Modeling the Benthos-Water Column Exchange of Hydrophobic Chemicals*. Ed. P. Gschwend, Technical Report EPA/600/S3-86/044, EPA Environmental Research Laboratory, Washington, DC. pp. 1-9.

EPA (U.S. Environmental Protection Agency). 2002. Great Lakes Contaminated Sediments: In-situ Cap Design. Volume 3. U.S. Environmental Protection Agency. URL (Last updated 1 August 2002): http://www.epa.gov/glnpo/sediment/iscmain/three.html

Fisher, J.B. 1982. Effects of macrobenthos on the chemical diagenesis of freshwater sediments. In Chapter 4 *Animal-Sediment Relations: The Biogenic Alteration of Sediments*, P.L. McCall and M.J.S. Tevesz (Eds.), Plenum Press, New York, NY. pp. 177-218.

Fisher, J.B., P.L. McCall, W.J. Lick, and J.A. Robbins. 1980. Vertical mixing of lake sediments by tubificid oligochaetes. *Journal of Geophysical Research* 85:3997-4006.

Ford, J.B. 1962. The vertical distribution of larval chironomidae (Diptera) in the mud of a stream. *Hydrobiologia* 19:262-272.

Godbout, L. and H.B.N. Hynes. 1982. The three dimensional distribution of the fauna in a single riffle in a stream in Ontario. *Hydrobiologia* 97:87-96.

Green, A.S. and G.T. Chandler. 1994. Meiofaunal bioturbation effects on the partitioning of sediment-associated cadmium. *Journal of Experimental Marine Biology and Ecology* 180:59-70.

Gunnison, D., J.M. Brannon, T.C. Sturgis, and I. Smith, Jr. 1987. *Development of a Simplified Column Test for Evaluation of Thickness of Capping Material Required to Isolate Contaminated Dredged Material*. Miscellaneous Paper D-87-2, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS. NTIS No. AD A179 929.

Jernelov, A. 1970. Release of methyl mercury from sediments with layers containing inorganic mercury at different depths. *Limnology and Oceanography* 15:958-960.

Karickhoff, S.W. and K.R. Morris. 1985. Impact of tubificid oligochaetes on pollutant transport in bottom sediments. *Environmental Science and Technology* 19:51-56.

Kirchner, W.B. 1975. The effect of oxidized material on the vertical distribution of freshwater benthic fauna. *Freshwater Biology* 5:423-429.

Krantzberg, G. 1985. The influence of bioturbation on physical, chemical and biological parameters in aquatic environments: a review. *Environmental Pollution (Series A)* 39:99-122.

Krantzberg, G. and P.M. Stokes. 1984. Benthic macroinvertebrates modify copper and zinc partitioning in freshwater-sediment microcosms. *Canadian Journal of Fisheries and Aquatic Sciences* 42:1465-1473.

Krezoski, J.R., S.C. Mozley, and J.A. Robbins. 1978. Influence of benthic macroinvertebrates on mixing of profundal sediments in Southeastern Lake Huron. *Limnology and Oceanography* 23:1011-1016.

Krezoski, J.R., J.A. Robbins, and D. White 1984. Dual radiotracer measurement of zoobenthos-mediated solute and particle transport in freshwater sediments. *Journal of Geophysical Research* 89:7937-7947.

Larsson, P. 1983. Transport of ^{14}C-labelled PCB compounds from sediment to water and from water to air in laboratory model systems. *Water Research* 17:1317-1326.

Maridet, L., M. Philippe, J.G. Wasson, and J. Mathieu. 1996. Spatial and temporal distribution of macroinvertebrates and trophic variables within the bed sediment of three streams differing by their morphology and riparian vegetation. *Archiv fur Hydrobiologie* 136:41-64.

Martin, P., X. Boes, B. Goddeeris, and N. Fagel. 2005. A qualitative assessment of the influence of bioturbation in Lake Baikal sediments. *Global and Planetary Change* 46:87-99.

Matisoff, G. and X. Wang. 1998. Solute transport in sediments by freshwater infaunal bioirrigators. *Limnology and Oceanography* 43:1487-1499.

Matisoff, G., X. Wang, and P.L. McCall. 1999. Biological redistribution of lake sediments by tubificid oligochaetes: *Branchiura sowerbyi* and *Limnodrilus hoffmeisteri*/*Tubifex tubifex*. *Journal of Great Lakes Research* 25:205-219.

McCall, P.L. and J.B. Fisher. 1979. Effects of tubificid oligochaetes on physical and chemical properties of Lake Erie sediments. *Proceedings of the First International Symposium on Aquatic Oligochaete Biology*. Plenum Press, Sidney, British Columbia, Canada. pp. 253-317.

McCall, P.L. and M.J.S. Tevesz. 1982. The effects of benthos on physical properties of freshwater sediments. In Chapter 3, *Animal-Sediment Relations: The Biogenic Alteration of Sediments*, P.L. McCall and M.J.S. Tevesz (Eds.). Plenum Press, New York, NY. pp. 105-176.

Mermillod-Blondin, F., G. Nogaro, T. Datry, F. Malard, and J. Gibert. 2005. Do tubificid worms influence the fate of organic matter and pollutants in stormwater sediments? *Environmental Pollution* 134:57-69.

Merritt, R.W. and K.W. Cummins. 1996. *An Introduction to the Aquatic Insects of North America.* Third Edition. Kendall/Hunt Publishing, Dubuque, IA. 862 p.

Morris, D.L. and M.P. Brooker. 1979. The vertical distribution of macro-invertebrates in the substratum of the upper reaches of the River Wye, Wales. *Freshwater Biology* 9:573-583.

Piercey, E. 1980. The importance of macrobenthic fauna to pore water chemistry and benthic fluxes in the transition zone of the Potomac River estuary. *Potomac Estuary Study: 1980 Fiscal Year Annual Report*. U.S. Geological Survey. pp. 87-93.

Poole, W.C. and K.W. Stewart. 1976. The vertical distribution of macrobenthos within the substratum of the Brazos River, Texas. *Hydrobiologia* 50:151-160.

QEA (Quantitative Environmental Analysis, LLC). 2005. *Comments of the General Electric Company on the U.S. Environmental Protection Agency's Model Calibration: Modeling of PCB Contamination in the Housatonic River*. Prepared for General Electric Company, Pittsfield, MA, by Quantitative Environmental Analysis, LLC, Liverpool, NY. February 7, 2005.

Rhoads, D.C., and L.F. Boyer. 1982. The Effects of Marine Benthos on Physical Properties of Sediments: A Successional Perspective. In *Animal-Sediment Relations: The Biogenic Alteration of Sediments*. P.L. McCall and M.J.S. Tevesz, Eds, Plenum Press, New York, NY. pp. 3-52.

Rhoads, D.C. and D.A. Carey. 1997. *Capping Dredged Materials in the New York Bight: Evaluation of the Effects of Bioturbation*. Science Applications International Corporation Report No. 374, Report No. 39 of the New York Mud Dump Series. Submitted to the U.S. Army Engineer District, New York.

Robbins, J.A. 1982. Stratigraphic and dynamic effects of sediment reworking by Great Lakes zoobenthos. *Hydrobiologia* 92:611-622.

Schaffner, L.C., R.M. Dickhut, S. Mitra, P.W. Lay, and C. Brouwer-Riel. 1997. Effects of physical chemistry and bioturbation by estuarine macrofauna on the transport of hydrophobic organic contaminants in the benthos. *Environmental Science and Technology* 21:3120-3125.

Shull, D.H. 2001. Transition-matrix model of bioturbation and radionuclide diagenesis. *Limnology and Oceanography* 46:905-916.

Swift, D.J.P., J.K. Stull, A.W. Niedoroda, C.W. Reed, and G.T.F. Wong. 1996. Contaminant dispersal on the Palos Verdes continental margin. II. Estimates of the biodiffusion coefficient DB from composition of the benthic infaunal community. *Sci. Tot. Environ*. 179:91-107.

Tevesz, M.J.S., P.L. McCall and F.M. Soster. 1980. The effects of size-selective feeding by oligochaetes on the physical properties of river sediments. *Journal of Sedimentary Petrology* 50:561-568.

Thomson, J., L. Brown, S. Nixon, G.T. Cook, and A.B. MacKenzie. 2000. Bioturbation and Holocene sediment accumulation fluxes in the north-east Atlantic Ocean (benthic boundary layer experiment sites). *Marine Geology* 169:21-39.

Thibodeaux, L.J. and V.J. Bierman. 2003. The bioturbation-driven chemical release process. *Environmental Science and Technology* 37:252A-258A.

van de Bund, W.J. and D. Groenendijk. 1994. Seasonal dynamics and burrowing of littoral chironomid larvae in relation to competition and predation. 1994. *Archiv fur Hydrobiologie* 132:213-225.

WESTON (Weston Solutions, Inc.). 2004a. *Ecological Risk Assessment for General Electric (GE)/Housatonic River Site Rest of River*. Environmental Remediation Contract, General Electric (GE)/Housatonic River Project, Pittsfield, MA. Contract No. DACW33-00-D-0006.

Task Order No. 0003. DCN: GE-100504-ACJS. Prepared for U.S. Army Corps of Engineers (New England District, Concord, MA) and U.S. Environmental Protection Agency (New England Region, Boston, MA) by Weston Solutions, Inc., West Chester, PA. November 12, 2004.

WESTON (Weston Solutions, Inc.). 2004b. *Model Calibration: Modeling Study of PCB Contamination in the Housatonic River,* Volume 4, *Appendix C, Bioaccumulation Model Calibration*. Environmental Remediation Contract, General Electric (GE)/Housatonic River Project, Pittsfield, MA. Contract No. DACW33-00-D-0006. Task Order No. 0003. DCN: GE-122304-ACMG. Prepared for U.S. Army Corps of Engineers (New England District, Concord, MA) and U.S. Environmental Protection Agency (New England Region, Boston, MA) by Weston Solutions, Inc., West Chester, PA. December 2004.

Wheatcroft, R.A. 1992. Experimental tests for particle size-dependent bioturbation in the deep ocean. *Limnology and Oceanography* 37:90-104.

Wheatcroft, R.A., I. Olmez, and F.X. Pink. 1994. Particle bioturbation in Massachusetts Bay: Preliminary results using a new deliberate tracer technique. *Journal of Marine Research* 52:1129-1150.

White, D.S., P.C. Klahr, and J.A. Robbins. 1987. Effects of temperature and density on sediment reworking by *Stylodrilus heringianus* (Oligochaeta: Lumbriculidae). *Journal of Great Lakes Research* 13:147-156.

Zwarts, L. and J. Wanink. 1989. Siphon size and burying depth in deposit- and suspension-feeding benthic bivalves. *Marine Biology* 100:227-240.

Table 1

Summary of Bioturbation Literature Review

Author(s)	Date	Description	Type of Organism/Species	Sediment Characteristics	Physical Environment	Bioturbation Depth Information	Notes	Most Relevant Habitat*	Relevance to PSA	Rationale
Clarke et al.	2001	Technical note presenting guidance on estimation of bioturbation profiles, depths, and process rates in relation to subaqueous sediment cap design	General	General	General	Freshwater silts/clays: 0-10 cm fully mixed; 10-30 cm partial mixing; freshwater sands: 0-10 cm fully mixed; 10-20 cm partially mixed	Relevant for all categories; mixing depths based on a literature review completed as part of the study.	1, 2, and 3	High	Sufficient data on mixing and sediment characteristics
Davis	1974b	Study of stratigraphic effects of tubificids in profundal lake sediment	Tubificid worms	Messalonskee Lake sediment; TOC 13-19% (dry weight basis)	Microcosm: freshwater lake sediment, total sediment depth >50 cm	Feeding primarily at 2-5 cm below surface; displacement of tracers predominantly in upper 8 cm and exclusively in upper 15 cm		3	High	Sufficient data on mixing and sediment characteristics
Fisher et al.	1980	Study of vertical mixing of lake sediment by tubificid oligochaetes	Oligochaete (*Tubifex tubifex*)	Silt-clay from Lake Erie	Microcosm: freshwater lake sediment, total sediment depth 13-15 cm	Complete mixing up to 7 cm; some mixing up to 9 cm		2 or 3	High	Sufficient data on mixing and sediment characteristics
Krezoski et al.	1984	Study of maximum depth of radio-labeled compound redistribution	Oligochaete (*Stylodrilus heringianus*), amphipod (*Pontoporeia*)	Sandy silt from southern Lake Michigan	Microcosm: freshwater, 20 cm sediment depth	4-5 cm for oligochaetes, 1.5 cm for amphipods	Study also includes sediment transport rates.	2	High	Sufficient data on mixing and sediment characteristics
Krezoski et al.	1978	Study of influence of benthos on mixing of profundal sediment in Lake Huron	Amphipods, Oligochaetes, Pelecypoda (some Mysidae, Diptera)	Sandy silt from Lake Huron	In situ: Lake Huron	3-6 cm (greater mixing depth with greater oligochaete abundance)		2	High	Sufficient data on mixing and sediment characteristics
Larsson	1983	Study of transport of PCBs from sediment to water and from water to air	Chironomids (*Chironomus plumosus*-type) and tubificids (*Tubifex tubifex*)	Sandy sediment from Lake Havgardssjon (southern Sweden)	Microcosm: freshwater, sediment depth 11 cm	PCB mixing down to 3 cm		1	High	Sufficient data on mixing and sediment characteristics
McCall and Fisher	1979	Study of effects of oligochaetes on physical and chemical properties of lake sediment	Tubificid oligochaetes (*Tubifex tubifex*)	Lake Erie mud (clay content 21% to 65%; TOC 6% to 9%)	Microcosm: freshwater, Lake Erie sediment (total depth 10 cm)	Well mixed at 5-10 cm	Sediment depth in microcosm may not have been sufficient to adequately characterize mixing depths.	3	High	Sufficient data on mixing and sediment characteristics
Campbell et al.	1988	Review of information on biologically available metals in sediment. Includes some general information on bioturbation	General	General	General	Well mixed up to 6 cm; no mixing below 20 cm.		General	Moderate	Sufficient data on mixing, insufficient information on sediment characteristics
Davis	1974a	Study of effect of tubificids on redox potential and pH in profundal lake sediment	Tubificids (*Limnodrilus hoffmeisteri*, *L. claparedeianus*)	Messalonskee Lake (Maine) sediment: noncalcareous copropelic gyttja, TOC 13-19%	Microcosm: freshwater, sediment depth 18 cm	Not directly provided	Tubificids influenced redox and pH in upper 15 cm of sediment.	3	Moderate	Information on influence of benthos on sediment, mixing depths could not be identified

Table 1

Summary of Bioturbation Literature Review (Continued)

Author(s)	Date	Description	Type of Organism/Species	Sediment Characteristics	Physical Environment	Bioturbation Depth Information	Notes	Most Relevant Habitat*	Relevance to PSA	Rationale
EPA	1987	Development of mathematical descriptions of processes controlling the exchange of chemicals between sediment and overlying water; depth of re-working provided for various species from other sources	Various	Not given	Not given	Marine gastropods: 2 mm; estuarine bivalves: 2 cm; freshwater oligochaetes: 4-6 cm	Mixing depths cited in the paper are based on data from other sources.	General	Moderate	Sufficient data on mixing, insufficient information on sediment characteristics
EPA	2002	Overview of considerations for in situ cap design	General	Sand, silt, clay	Great Lakes - freshwater	Approximately 5-10 cm in sand; greater in silts and clays	Burrowing up to 10-40 cm expected for a silt or clay; depth of bioturbation by marine benthos generally greater than freshwater benthos.	General	Moderate	General information regarding sands provided; insufficient information regarding physical or biological environment
Fisher	1982	Review of effects of macrobenthos on chemical changes in sediment occurring during and after burial	General	General	General freshwater	Amphipods, sphaerid clams: 1-3 cm; tubificid oligochaetes, unionid clams: ~10 cm		2 or 3	Moderate	Sufficient data on mixing, insufficient information on sediment characteristics
Jernelov	1970	Study of release of methylmercury from sediment	Tubificidae and Anodonta (mussels)	Eutrophic lake sediment; grain size and TOC not provided	Microcosm: freshwater, sediment depth 20 cm	Not directly provided	Release from up to 3 cm for Tubificidae, up to 10 cm for Anodonta; extent of mixing not quantified; sediment physical characteristics not provided.	Likely 2 or 3	Moderate	Sufficient data on mixing, insufficient information on sediment characteristics
Karickhoff and Morris	1985	Study of impact of tubificid oligochaetes on pollutant transport in sediment; focus not on mixing depth, limited microcosm sediment depth	Tubificid oligochaetes (*Limnodrilus hoffmeisteri*, *Tubifex tubifex*)	Sand (50-80%); TOC 0.5-1.5%	Microcosm: freshwater stream sediment, total depth 2-5 cm	At least 3 cm	Based on interpretation of data presented graphically in the paper.	1	Moderate	Data on mixing inferred from the information provided in the paper
Krantzberg	1985	General review of the influence of bioturbation on physical, chemical, and biological parameters in aquatic environments	General benthic invertebrates; oligochaetes and amphipods considered important for bioturbation	Various	Various	3-20 cm		General	Moderate	Sufficient data on mixing, insufficient information on sediment characteristics
Krantzberg and Stokes	1984	Study of effect of benthos on copper and zinc partitioning in sediment; focus on geo-chemical changes and release of metals to water column	Chub Lake: Chironimidae, Tanypodinae, Chaoborus Lohi Lake: Chironomidae, Chaoborus, Tubificidae Port Credit: Tubificidae	Freshwater lake sediment, no details on grain size or TOC	Microcosm studies using sediment from Chub Lake, Lohi Lake, and Port Credit; total sediment depth 10 cm	Not directly provided	Data indicate possible greater influence of benthos in upper 5 cm compared to 5-10 cm depth (microcosm sediment depth = 10 cm).	General	Moderate	Sufficient data on mixing, insufficient information on sediment characteristics
Matisoff and Wang	1998	Study of solute transport in sediment by infaunal bio-irrigators	Chironomids (*Coelotanypus* sp. and *Chironomus plumosus*) and mayfly larvae (*Xexagenia limbata*)	Lake Erie sediment: silt and clay	Microcosm: 20-25 cm depth	Not directly provided	Solute transport depth: *Coelotanypus* sp: 3.0 cm; *Chironomus plumosus*: 5-9 cm; mayfly: 5 cm.	2 or 3	Moderate	Focus on influence of benthos on solute transport and not sediment mixing

Table 1

Summary of Bioturbation Literature Review (Continued)

Author(s)	Date	Description	Type of Organism/Species	Sediment Characteristics	Physical Environment	Bioturbation Depth Information	Notes	Most Relevant Habitat*	Relevance to PSA	Rationale
Matisoff et al.	1999	Study of biological redistribution of lake sediment by tubificid oligochaetes	Oligochaetes: *Branchiura sowerbyi*, harbour worms (*Limnodrilus hoffmeisteri*, *Tubifex tubifex*)	Lake Erie mud	Microcosm: freshwater, sediment depth 10-15 cm	Harbour worms: well mixed to approx. 5 cm, some mixing to at least 10 cm; *B. sowerbyi*: well mixed to approx. 13 cm	Based on interpretation of data presented graphically in the paper.	2 or 3	Moderate	Sufficient data on mixing, insufficient information on sediment characteristics
McCall and Tevesz	1982	General review of effects of benthos on freshwater sediment physical properties	General	General	General	Mixing depth varies but generally well mixed up to 8 cm or less	Oligochaete feeding depth varies up to 8 cm, depending on species and density.	General	Moderate	Sufficient data on mixing, insufficient information on sediment characteristics
Mermillod-Blondin et al.	2005	Study of influence of tubificids on fate of organic matter and pollutants in stormwater sediment; focus on biogeochemical processes	Tubificid worms (*Tubifex tubifex*)	Sediment from bed of a stormwater infiltration basin, no information on grain size or TOC	Microcosm: total sediment depth 10 cm	Not directly provided	Tubificids appear to influence primarily the upper 5 cm of sediment.	General	Moderate	Sufficient data on mixing, insufficient information on sediment characteristics
Thibodeaux and Bierman	2003	General review of the process of bioturbation-driven release of chemicals from the sediment bed	General	General	General	Not directly provided	Bioavailable surface layer typically around 10 cm depth.	General	Moderate	Sufficient data on mixing, insufficient information on sediment characteristics
Bishop	1973	Study of vertical distribution of benthos in a Malaysian stream; no details on mixing	General (primarily chironomids)	Some fine sediment mixed with coarser sands and gravels	In situ, freshwater, moderately flowing riffle area	Not directly provided	Highest percentage of organisms found at 0-10 cm depth but significant proportions found up to 40 cm.	-	Low	Vertical distribution data only, no information on effects of benthos on sediment
Boudreau	1998	Development of a resource-feedback model of bioturbation that predicts the existence of a finite-depth bioturbated zone with a universal mean mixing depth of 9.7 cm, which is consistent with a previously defined mixing depth of 9.8 cm ± 4.5 cm, based on empirical data from marine environments	General	Marine sediment (general)	Marine sediment (general).	9.7 cm		-	Low	Marine
Boudreau	1994	Evaluation of relationship among mixing coefficient, mixing depth, and burial velocity using data from coastal, shelf, slope, and deep-sea environments	General	Marine sediment (general)	Marine sediment (general)	9.8 ± 4.5 cm		-	Low	Marine
Brannon et al.	1985	Study of cap depth required to isolate contaminated dredged material from biota and overlying water	Clam (*Rangia cuneata*), Polychaete (*Nereis virens*)	3 types of capping materials: sand, silt, and clay	Microcosm: marine, varying sediment and cap depths	Not directly provided	A cap of 50 cm was sufficient to prevent chemical transfer from contaminated sediment to overlying water; a cap of 5 cm was not sufficient.	-	Low	Marine
Cunningham et al.	1999	Focus on pyrene loss in upper 15 mm sediment due to flux, degradation, and/or enhanced microbial processes associated with bioturbation	Oligochaete (*Limnodrilus cervix*)	Campus Lake sediment: silty loam; TOC 1.12%	Microcosm: freshwater, 3.5 cm depth of sediment	Not directly provided	Effects on sediment geo-chemistry observed throughout upper 1.5 cm (no data for deeper sediment).	-	Low	Insufficient study depth: focus on geochemistry of upper 1.5 cm

Table 1

Summary of Bioturbation Literature Review (Continued)

Author(s)	Date	Description	Type of Organism/Species	Sediment Characteristics	Physical Environment	Bioturbation Depth Information	Notes	Most Relevant Habitat*	Relevance to PSA	Rationale
Danielopol	1976	Study of distribution of invertebrates in river sediment; no information on mixing; coarse sample intervals (i.e., large depth range for samples)	Various	Various (freshwater)	Danube River	Not directly provided	Highest densities in the upper 0.5 m.	-	Low	Vertical distribution data only, no information on effects of benthos on sediment
Davis	1983	Study of effect of polychaetes on sediment-water interface and exchange of copper between sediment and sea water	Polychaete (*Nephtys incisa*)	Fine-grained, high silt-clay; 10-11% TOC	Microcosm: marine	Total bioturbation zone 20-30 cm		-	Low	Marine
Ford	1962	Study of distribution of invertebrates; no information on mixing	Various	Mud from a small stream	Stream	Not directly provided	Chironomids: 79% in upper 2.5 cm, 98% in upper 5 cm; oligochaetes found throughout total sample depth (upper 20 cm); other less abundant invertebrates found up to 15 cm depth.	-	Low	Vertical distribution data only, no information on effects of benthos on sediment
Godbout and Hynes	1982	Study of distribution of invertebrates in a creek	Various	Cobbles with pebbles, gravel, and sand	Uniform riffle, Salem Creek (Ontario)	Not directly provided	Most invertebrates present in upper 25 cm, highest abundance in 5-15 cm sample.	-	Low	Vertical distribution data only, no information on effects of benthos on sediment
Green and Chandler	1994	Study of meiofaunal bioturbation effects on partitioning of sediment-associated cadmium	Harpacticoid copepod (*Amphiascus tenuiremis*), Foreminiferan (*Ammonia beccarii*)	<63 µm, TOC 3.8%	Microcosm: sediment depth 1.2 cm; salinity 30.4 ppt	Foraminifora: 0-3 mm well mixed; 3-5 mm some mixing; copepods: 0-5 mm well mixed; 5-1.2 mm some mixing		-	Low	Marine
Gunnison et al.	1987	Development and evaluation of a small-scale predictive test for identifying minimum cap thickness for contaminated sediment	Various	Various types of capping materials	Microcosm: marine, varying sediment and cap depths	Not directly provided	A cap depth of 22 to 50 cm was required for isolating contaminated sediment from benthos and overlying water.	-	Low	Marine
Kirchner	1975	Study of effect of oxidized material on vertical distribution of benthos. Focus on vertical distribution of benthic invertebrates	Ostracoda, Nematoda, Oligochaeta, Harpacticoida, Cyclopoida, Chironomidae	Fine silt (primarily dolomite)	Char Lake, Northwest Territory	Not directly provided	Biomass predominantly in upper 3-4 cm.	-	Low	Vertical distribution data only, no information on effects of benthos on sediment
Maridet et al.	1996	No information on mixing but detailed information on vertical distribution of organisms	Primarily Chironomidae, Elmidae, and Leuctridae	Stream sediment; no details on grain size or TOC	Mountain stream sediment (France), in situ	Not directly provided	70-96% of individuals found in upper 15 cm.	-	Low	Vertical distribution data only, no information on effects of benthos on sediment

Table 1

Summary of Bioturbation Literature Review (Continued)

Author(s)	Date	Description	Type of Organism/Species	Sediment Characteristics	Physical Environment	Bioturbation Depth Information	Notes	Most Relevant Habitat*	Relevance to PSA	Rationale
Martin et al.	2005	Qualitative assessment of bioturbation influence on Lake Baikal sediment	Various (most abundant were oligochaetes, nematodes, copepods)	Mud	Lake Baikal (water depths 128 m to 623 m) - freshwater	6 cm at water depth of 128 m, with benthos penetration to 15 cm; thinner depths observed at deeper sites	Depth of mixing decreased with increasing water depth.	-	Low	Deep lake environment
Morris and Brooker	1979	Study of vertical distribution of macroinvertebrates	Various	Not given	River Wye, Wales	Not directly provided	82% of organisms found in upper 22 cm.	-	Low	Vertical distribution data only, no information on effects of benthos on sediment
Piercey	1980	Study of effect of benthic community on pore water chemistry	Various	Not given	Potomac River estuary, fluctuating salinity	Not directly provided	Pore water in top 10 cm of well-burrowed sediment was well mixed; poorly burrowed sediment had a diffusion gradient.	-	Low	Marine
Poole and Stewart	1976	Study of abundance of benthos at various depths in Brazos River sediment; no information on mixing	Various	Coarse gravel	Brazos River	Not directly provided	Macrobenthos predominantly present in upper 20 cm, but significant numbers found up to 40 cm depth.	-	Low	Vertical distribution data only, no information on effects of benthos on sediment
Rhoads and Carey	1997	Role of bioturbation in the context of cap design in estuarine and coastal settings	Various	Various types of capping materials	Coastal and estuarine	Not directly provided	Colonization of organisms on capped materials results in deeper feeding after several years (approximately 10 cm biologically mixed depth). Late stages of colonization may develop large organisms that feed to depths close to 30 cm.	-	Low	Marine
Schaffner et al.	1997	Effects of physical chemistry and bioturbation by estuarine macrofauna on the transport of hydrophobic organic contaminants in benthos	Not given	Marine (>18 ppt salinity) Chesapeake Bay sediment; 50% sand, 35% silt, 15% clay; <1% TOC	Microcosm: sediment depth 14 cm, saline (~20 ppt)	Contaminant burial to maximum 2 cm after 56 days	-	-	Low	Marine
Shull	2001	Development of a model of bioturbation and comparison with field data	Bivalve (*Nucula annulata*, *Yoldia limatula*), polychaete (*Mediomastus ambiseta*, *Polycirrus eximius*, *Sabaco elongatus*, *Macroclymene zonalis*)	Narragansett Bay sediment: TOC 0.9%-1.2%	In situ, marine	Well-mixed to 5 cm; some mixing 5 cm-15 cm		-	Low	Marine
Swift et al.	1996	Development of time- and space-dependent biodiffusion coefficients for the sea bed, based on the composition and distribution of the benthic community	Various	Not documented	60-m isobath off the Palo Verdes shelf, California	Fully mixed zones ranged from under 5 cm to almost 30 cm	Mixing rates provided.	-	Low	Marine environment

Table 1

Summary of Bioturbation Literature Review (Continued)

Author(s)	Date	Description	Type of Organism/Species	Sediment Characteristics	Physical Environment	Bioturbation Depth Information	Notes	Most Relevant Habitat*	Relevance to PSA	Rationale
Tevesz et al.	1980	Describes stratification of sediment by oligochaete feeding	Oligochaete	Muddy sand from Vermillon River (OH)		Not directly provided	Layer formation in upper 1.5 cm.	-	Low	
Thomson et al.	2000	Study of bioturbation and Holocene sediment accumulation fluxes in the northeast Atlantic Ocean	Various	Carbonate materials overlying clays at depths greater than 15 cm	Northeast Atlantic Ocean	Range of 8 to 17 cm	-	-	Low	Marine, deep ocean
Van de Bund and Groenedijk	1994	Study of chironomid larvae distribution	Chironomid larvae	Sand	Oligo-mesotrophic lake	Not directly provided	Larvae found almost exclusively (>99%) in upper 5 cm.	-	Low	Vertical distribution data only, no information on effects of benthos on sediment
Wheatcroft	1992	Study of bioturbation in deep ocean and the influence of particle size	Various	Various	Deep ocean	Not directly provided		-	Low	Marine
Wheatcroft et al.	1994	Study of particle bioturbation in Massachusetts Bay using tracers	Various	Massachusetts Bay sediment, sand, and silt	In situ microcosms: marine, one with silt and one with sand	Not directly provided	Spring: average vertical penetration 4-6 cm, max. 15 cm; sand average vertical penetration <2 cm, max. 6 cm; fall: sand and silt both had peaks at 10 cm and 20 cm.	-	Low	Marine
White et al.	1987	Study of sediment reworking by oligochaetes over a range of temperatures and organism densities	Oligochaete (*Stylodrilus heringianus*)	Lacustrine muds	Mesocosm	Not directly provided	Depth to which organism fed was positively related to worm density. Report cites other studies (Robbins, 1982) in which constant mixing of as much as the upper 15 cm of sediment occurs in Laurentian Great Lakes.	-	Low	No information on well mixed layers; study emphasized reworking rates
Zwarts and Wanink	1989	Study of relationship between burying depth and shell size for bivalves; no information provided on mixing	Bivalves (*Mya arenaria*, *Cerastoderma edule*, *Scrobicularia plana*, *Macoma balthica*)	Mud; clay content 6.5%	Intertidal mudflats of the Dutch Wadden Sea, marine	Not directly provided	Approximate maximal burying depths: *M. arenaria* 25 cm, *C. edule* 2 cm, *M. balthica* 8 cm, *S. plana* 15 cm.	-	Low	Marine

1 * See Table 2 for description of habitats; codes assigned only for studies with high or moderate relevance to PSA.

Table 2

Broad Habitat Types in Housatonic River PSA Sediment

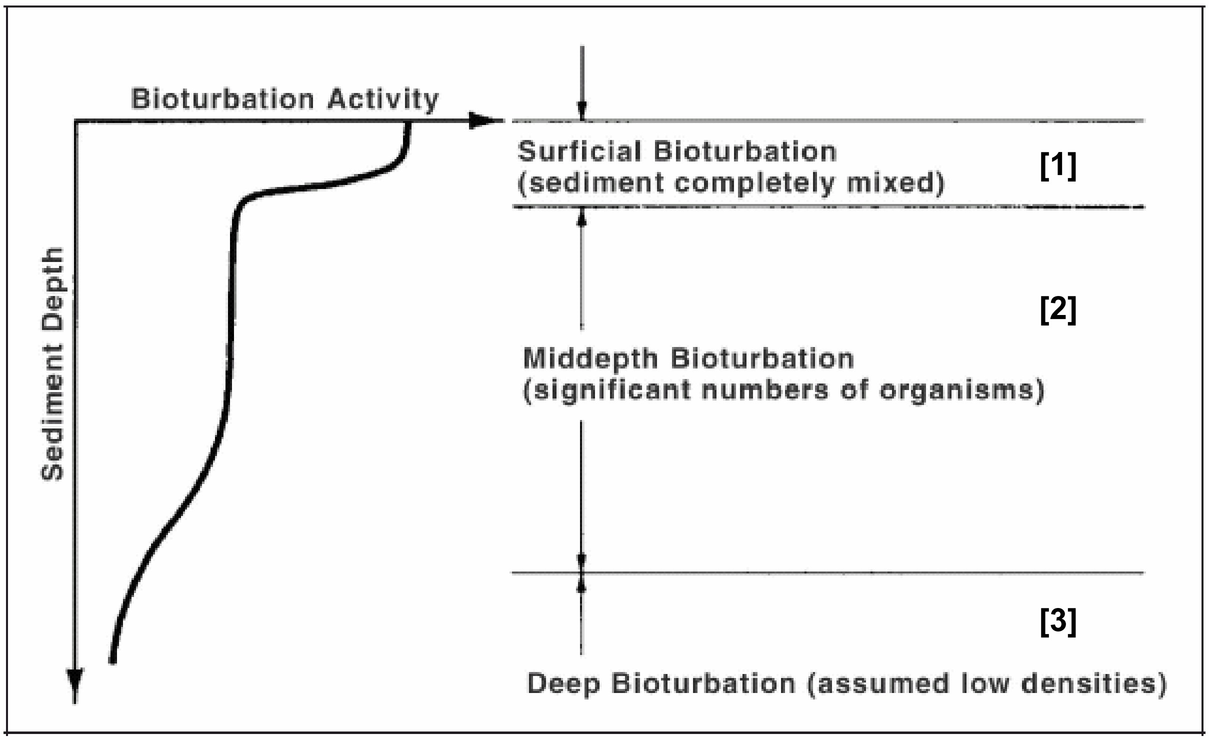
Habitat		Dominant Sediment Type	Typical TOC	Dominant Organism Types	Relative Biota Density
1	Upstream main channel (Reach 5A)	Medium sand	1%	Many non-burrowing chironomids (e.g., *Tanytarsus* sp.), some burrowing chironomids (e.g., *Saetheria* sp.), tubificid oligochaetes (e.g., *Limnodrilus* sp.), sphaeriid clams, some mayflies, some caddisflies.	Low
2	Downstream (Reaches 5B, 5C, Woods Pond)	Silt, some fine sand	5%-6%	Many sediment burrowing chironomids (e.g., *Dicrotendipes* sp., *Einfeldia* sp.), some non-burrowing chironomids (e.g., *Polypedilum* sp.), tubificid oligochaetes (e.g., *Limnodrilus* sp.), gastropods (e.g., Hydrobiidae), sphaerid clams, amphipods, damselflies.	High
3	Backwaters (Reach 5D)	Mainly fines (silt and clay)	10-15%	Expected to be similar to Area 2, perhaps with increased proportion of deposit feeders.	Not quantified; inferred to be high

Table 3

Summary of Biological Mixing Depths in PSA Sediment

Habitat		Biologically Mixed Depth Interval *	Biologically Influenced Depth Interval *
1	Upstream main channel (Reach 5A)	0 – 4 cm	4 – 10 cm
2	Downstream (Reaches 5B, 5C, Woods Pond)	0 – 8 cm	8 – 15 cm
3	Backwaters (Reach 5D)	0 – 10 cm	10 – 20 cm

* The depths shown represent depth intervals below the sediment surface. The total bioturbation depth is represented by the bottom of the biologically influenced depth interval (e.g., 10 cm for Reach 5A). Definitions for terms are provided in Section 1.2.



Notes. [1] Biologically mixed layer – high densities of infaunal organisms completely rework the sediment matrix.

[2] Biologically influenced layer – mid-depth zone, in which infaunal densities are lower and rates of sediment reworking decrease with depth.

[3] Deep bioturbation zone – bioturbation potential is negligible to low.

Figure extracted from Clarke et al. (2001).

Figure 1 Concept of Bioturbation "Zones" that Correspond to Intensities and Vertical Distribution of Organisms

ATTACHMENT 2

KOLMOGOROV-SMIRNOV APPROACH TO EVALUATION OF UNCERTAINTY FOR EFDC

ATTACHMENT 2

KOLMOGOROV-SMIRNOV APPROACH TO EVALUATION OF UNCERTAINTY FOR EFDC

1. INTRODUCTION

As discussed in the response to General Issue 11, this attachment provides technical details for the Kolmogorov-Smirnov confidence limit approach that will be used to evaluate the uncertainty of model output(s) for EFDC and which will also be used in the evaluation of uncertainty propagated through the linked models (see response to General Issue 11).

2. KOLMOGOROV-SMIRNOV CONFIDENCE LIMIT APPROACH

Kolmogorov-Smirnov confidence limits (Sokal and Rohlf, 1994) provide uncertainty bounds at a specified probability level (α) for an entire empirical histogram or frequency distribution. This method can be used to evaluate uncertainty for any model parameter and for any grouping of model cells. The simulated values for the parameter of interest as output by the model, e.g., total PCB (tPCB) concentration, are the components of the frequency distribution. The bounds are constructed by computing, at each observed value of the empirical variable, the observed cumulative frequency plus and minus the critical value of the Kolmogorov-Smirnov function (D_α). This provides $100(1-\alpha)\%$ confidence limits around the simulated cumulative distribution function. This method requires specification of both the range (outside of which the distribution is truncated) and the confidence level to be used. In practice, a confidence level of $\alpha = 0.05$ is typical, and the range can be derived from the data. For example, a reasonable lower bound might be the analytical detection limit for the parameter of interest, and a reasonable upper bound might be the highest measured value for the parameter.

Figure 1 illustrates an empirical cumulative frequency distribution constructed from a sample of six points (0.2, 0.5, 0.6, 0.7, 0.75, and 0.8) for a variable constrained between 0 and 1. The empirical distribution function histogram is shown in gray, and the 95% Kolmogorov-Smirnov limits are shown in black. Of the time that confidence limits are generated by random sampling from the distribution (95%), they will bound completely the true cumulative frequency distribution. The bounds would be tighter if there were more data points, or if a lower confidence level were used.

EPA proposes to use the Kolmogorov-Smirnov method for evaluating uncertainty in EFDC simulations on the same spatial scale as the data are provided by EFDC to FCM. In such an application, the uncertainty in simulated results for the parameters of interest (tPCBs in sediment and particulate organic matter [POM], respectively, are the principal parameters to be examined because they are the parameters output from EFDC to which FCM is most sensitive) will be evaluated for the modeling subreaches used by FCM, i.e., 5A, 5B, 5C, 5D, and 6. At a minimum, the uncertainty analysis will be conducted for results at the end of the Phase 2 Calibration period (1990 to 2000), but may also be evaluated at the conclusion of any daily time step, which is the frequency at which EFDC passes data to FCM, during the entire validation period (1979 to 2004).

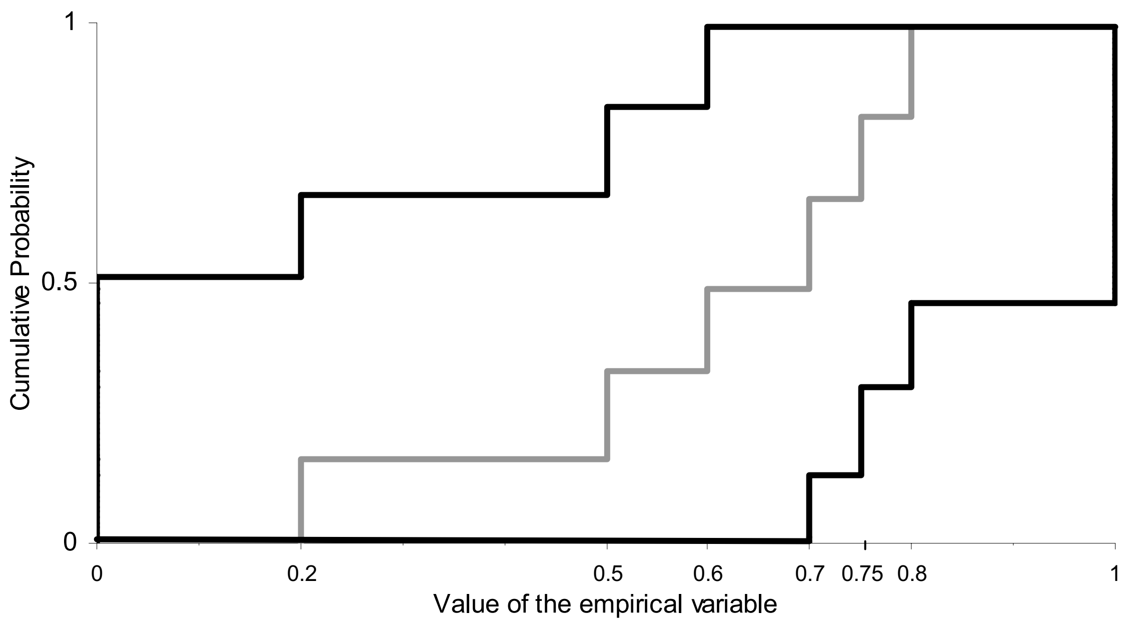


Figure 1 Kolmogorov-Smirnov 95% Confidence Bounds on an Empirical Distribution Made up of Eight Values (0.2, 0.5, 0.6, 0.7, 0.75, and 0.8) Assuming the Distribution Ranges from 0 to 1

3. REFERENCES

Sokal, R.R. and F.J. Rohlf. 1994. *Biometry.* 3rd edition. Freeman and Company, New York, NY.