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 VALIDATION:

MODELING STUDY OF PCB CONTAMINATION IN THE HOUSATONIC RIVER

DCN: GE-111006-ADII

November 2006

Environmental Remediation Contract
GE/Housatonic River Project
Pittsfield, Massachusetts

Contract No. DACW33-00-D-0006

Task Order 0003

06P-1130-1

WESTON SOLUTIONS®

RESPONSIVENESS SUMMARY TO THE PEER REVIEW OF MODEL VALIDATION:
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-111006-ADII

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

November 2006

Work Order No. 20123.001.096.0683

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, Rooni Mathew, Mark Velleux
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**

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

REFERENCES

TABLE OF CONTENTS
(Continued)

APPENDIX A PEER REVIEWER COMMENTS

LIST OF ACRONYMS

CFD	computational fluid dynamics
cm/s	centimeters per second
CMS	Corrective Measures Study
DRM	Document Review Meeting
EDS	energy-dispersive system
EFDC	Environmental Fluid Dynamics Code
EPA	U.S. Environmental Protection Agency
EqP	equilibrium partitioning
FCM	food chain model
FMD	Final Model Documentation
FN	Floodplain Number
GE	General Electric Company
HOCs	hydrophobic organic chemicals
HSPF	Hydrologic Simulation Program - FORTRAN
kg/year	kilograms per year
KS	Kolmogorov-Smirnov
MB	Model Bias
MCR	Model Calibration Report
MFD	Modeling Framework Design
mm/year	millimeters per year
MVR	Model Validation Report
NRCSE	National Research Center for Statistics and the Environment
PCB	polychlorinated biphenyl
PRC	Peer Review Committee
PSA	Primary Study Area
QAPP	Quality Assurance Project Plan
RFI	RCRA Facility Investigation
ROR	Rest of River
RSM	Response Surface Modeling
SEM	scanning electron microscopy
SNL	Sandia National Laboratories
SPI	sediment profile imaging
SRA	SRA International, Inc.

LIST OF ACRONYMS (CONTINUED)

TSS	total suspended solids
UV-B	ultraviolet
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 Validation Report for the GE/Housatonic River Site Rest of River released in March 2006. The review was conducted by seven professionals in disciplines related to numerical modeling of aquatic and riverine systems. This document, referred to herein as the Model Validation Responsiveness Summary, has been prepared as part of EPA's obligations under Appendix J of the comprehensive agreement relating to the cleanup of the General Electric Company (GE) Pittsfield, MA facility, certain off-facility 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 stipulates that the model 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 45-day public comment period provided the opportunity for the public and GE to submit written comments on the Model Validation Report for consideration by the Peer Review Panel, within the context of the Peer Review Charge. On June 28 and 29, 2006, the Model Validation Peer Review Panel ("Reviewers") met at a public forum in Lenox, MA, to review and discuss the Model Validation Report within the framework of the Charge. During this meeting, 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. (SRA), of Arlington, VA.

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. During the Peer Review meeting, many of the Reviewers provided some of the same comments on the Model Validation Report; therefore, they submitted similar written comments on these topics. Conversely, at many times Reviewers had differing views on topics; 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 Validation Responsiveness Summary, EPA organized this document so that responses to general topics are grouped together, followed by responses to other (miscellaneous) comments.

The responses to the General Topics address subjects that were raised by a number of Reviewers and/or had broad implications for the model validation and for the modeling study in general. EPA has identified 12 General Topics and one Other (Miscellaneous) Topic and has provided a Summary of Topic for each to frame the technical basis for the issue and to provide an indication of how often the topic was noted by the Reviewers. Typically, each Summary of Topic is followed by EPA's response to the General Topic. Most of the responses to the General Topics reference the Final Model Documentation (FMD) Report to avoid unnecessary repetition.

Appendix A presents the Peer Reviewer Comments and indicates the classification of these comments in relation to the General Topics.

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

This Responsiveness Summary provides EPA's formal response to the final written Peer Review comments and, together with the issuance of the FMD Report and meetings/discussions of the Model Working Group transferring the model code and input files, constitutes EPA's declaration of completion of the Model Validation Report Peer Review and fulfillment of the requirements of Paragraph 22.j of the Consent Decree.

The purpose of the FMD is to summarize the enormous effort resulting from 8 years of interpretation, analysis, and evaluation that occurred in support of the modeling study for the Housatonic River, incorporating input from GE and its consultants as part of the Model Working Group established under the Consent Decree, and in response to the Peer Reviews. The FMD finalizes the outcome of the modeling study and draws upon the information contained in the following documents:

- Draft Model Framework Design (MFD) document and Quality Assurance Project Plan (QAPP) (October 2000).
- Responsiveness Summary to the Peer Review of the MFD and QAPP (June 2002).
- Final MFD (April 2004).
- Model Calibration Report (MCR) (December 2004).
- Responsiveness Summary to the Peer Review of the MCR (January 2006).
- Model Validation Report (MVR) (March 2006).
- Responsiveness Summary to the Peer Review of the MVR (November 2006).

- Other materials prepared in response to the Peer Reviews that are included in the Peer Review record.
- The RCRA Facility Investigation (RFI) Report prepared by GE (September 2003).

The FMD is organized in roughly parallel construction to both the Model Calibration and Validation Reports to assist the reader in going back to the source document to obtain more detail regarding the subject being presented, if desired.

The comments received from Reviewers on the Model Validation Report were used by EPA in preparing this Responsiveness Summary and the FMD. In many cases the modeling team performed further evaluations of the data or model formulations based upon suggestions by the Reviewers. In some cases, these resulted in changes in the model framework. These efforts and resulting changes are discussed in this Responsiveness Summary and documented in the FMD.

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 the FMD; in such cases, the technical basis for EPA's position is provided.

RESPONSE TO GENERAL TOPICS AND SPECIFIC COMMENTS

1. GENERAL TOPIC 1: REPRESENTATION (PARAMETERIZATION) OF MODEL PROCESSES

1.1 BALANCE OF FATE PROCESSES

1.1.1 Comment Summary

Reviewer Name	Page Number/Line Number
Bohlen	Page 4, lines 23 to 33
	Page 8, line 38 to page 9, line 7
	Page 9, lines 8 to 23
	Page 14, lines 26 to 31
Endicott	Page 2, lines 20 to 21
Garcia	Page 1, lines 25 to 37
	Page 5, line 39 to page 6, line 2
Gobas	Page 2, lines 28 to 42
	Page 3, lines 34 to 37
	Page 5, lines 4 to 17
	Page 13, lines 9 to 25
Lick	Page 1, lines 19 to 28
	Page 3, line 16 to page 4, line 10
	Page 9, lines 14 to 43
List	Page 5, lines 31 to 34
	Page 5, lines 39 to 40

1.1.2 General Topic Summary

The model framework for the Housatonic River includes HSPF, EFDC, and FCM. To achieve the goals of the modeling study, the model framework must reasonably account for the relevant processes that control PCB fate, transport, and bioaccumulation in the Housatonic River. In the framework, EFDC is used to simulate hydrodynamics, sediment transport, and PCB fate. Within EFDC, multiple processes are represented and interact to influence the transport and fate of sediment and contaminants and occur across different spatial and temporal scales. The representation of the interactions between processes and the relative importance of the relevant processes (i.e., the balance of fate processes), at the relevant scales, is central to determining the ability of the framework to simulate future conditions.

The majority of Reviewers expressed opinions that processes were either not well understood or incorrectly (or not accurately) represented in the model. Reviewers often recommended that additional information and/or data be collected to address their concerns. In some cases specific examples were provided and are summarized, with EPA's response, in the following subsections of the discussion of this General Topic.

1.1.3 General Topic Response

EPA strongly disagrees with Reviewer opinion that relevant sediment transport and/or PCB fate processes are not reasonably represented in the model framework. The balance of fate processes refers to the relative importance of processes for phase distribution, erosion, deposition, and sediment-water mass transfer that give rise to PCB exposure concentrations in water and sediment.

It is important to recognize that the overwhelming majority of Reviewer comments regarding the relative balance of fate processes described in the model framework are founded on erroneous interpretations of the rate at which PCB concentrations in river sediment have changed over time and how individual transport processes affect that rate. The change in sediment PCB concentrations over time was initially described in the RFI (BBL and QEA, 2003), and EPA documents such as Responses to Questions from the Reviewers, and is now more fully quantified in Appendix A.1 of the Final Model Documentation (FMD) Report. Specifically, these analyses indicate that the trend (or decline) in sediment PCB concentrations in Reaches 5 and 6 is not statistically different from zero (i.e., slow to no change over time).

Accurate quantification of spatial and temporal PCB concentration trends in river sediments is complex. Sediment sampling efforts for the Housatonic River were not designed to estimate PCB concentration trends over time. Differences in sediment PCB concentrations resulting from spatial heterogeneity, temporal variability, and analytical bias confound analysis and make clear identification of trends challenging (see FMD Appendix A.1 and General Topic 2).

The trend (or decline) in surface sediment PCB concentrations in Reaches 5 and 6 shows slow to no change over time. Model results are consistent with this observation. Specifically, the simulated rate of surface sediment PCB concentration decline in EFDC is slow and within the uncertainty envelope of the data. The analysis of spatial and temporal trends in PCB concentrations in fish tissue shows a slight decline over time in whole-body fish tissue PCB concentrations in Reaches 5 and 6. The decline in fish tissue concentrations was statistically significant for some of the species and/or subreaches and not for others. The magnitude of the decline was generally small in most cases, with a half-life on the order of decades. This is consistent with the results of the analysis of trends in sediment PCB concentrations. PCB exposure concentrations simulated in the water column and sediment and passed from EFDC to FCM result in the accurate simulation of the concentrations of PCBs in fish tissue, with a simulated rate of decline which is slow and within the uncertainty envelope of the data.

In contrast, the Reviewers assumed that there was a more rapid decline in sediment PCB concentrations over time and then incorrectly concluded that the rate of decline in PCB concentrations simulated by the model (which is in agreement with the data) was too slow. The Reviewers did not have a sound basis for their assumption, which led to their erroneous conclusions that the balance of fate processes was not correctly represented in the model framework. Therefore, many of the ensuing Reviewer comments regarding individual fate processes, the balance of fate processes, and the ability of the model to simulate future conditions, are logically flawed.

Implementation of an independent calibration strategy helped ensure that errors were minimized and that the balance of fate processes in the model framework is correctly represented. All models in the framework were calibrated independently. Information was passed forward from one model to the next in sequence, as presented in FMD Figure 1-11. Outputs from HSPF were used as inputs to EFDC and FCM. Further outputs from EFDC were used as inputs to FCM. Information was never passed backwards from FCM to EFDC or HSPF. No compensation was made in one model for a performance issue in another model. For example, FCM was not recalibrated to achieve acceptable model-data comparisons in fish tissue to compensate for any potential bias in flow, TSS, or PCB outputs from EFDC or HSPF.

In addition, a priori model performance measures were proposed in the Quality Assurance Project Plan, which was the subject of the first of the three Peer Reviews along with the Modeling Framework Design (MFD). These measures were subsequently adopted following Review comment. All three models, including EFDC, and more importantly, the linked FCM, produce results that are, in virtually all cases, within these performance measures. Achieving these performance measures, in combination with the successful implementation of this independent calibration strategy, demonstrates that model framework performance is adequate to meet the goal of the modeling study.

Many comments nonetheless focused on the balance of fate processes in EFDC. The Reviewers did not recognize that a fundamental attribute of the EFDC calibration strategy was that each transport and fate process in the model was independently examined. Specifically, the site data were reviewed to isolate conditions where sediment or PCB transport was controlled by a single, predominant process. To the greatest extent possible, values for each model process were determined using site-specific information. This calibration approach is particularly important to the overall model development effort because it ensured that: (1) each individual transport and fate process was correctly parameterized; and (2) the relative importance of each process is appropriately represented. In addition, calibration of the food chain model was performed independently of the calibration of EFDC. PCB exposure concentrations simulated in the water column and sediment and passed to FCM result in an accurate simulation of the concentrations of PCBs in fish tissue, which provides a further test of the adequacy and accuracy of the representation of the balance of fate processes in EFDC.

In other instances, the Reviewers have made recommendations for data collection efforts that fall well outside the spatial scales of intended model use or the timeframe for model development efforts. EPA has taken seriously Reviewer suggestions for data collection to improve model development. In fact, EPA undertook several data collection efforts in response to recommendations. However, it must be recognized that over the past 6 years, three Peer Reviews of the Model Framework, Calibration, and Validation have now been completed. While the Peer Review Charge for the Modeling Framework Design (2000) included specific questions asking Reviewers to recommend supporting data collection efforts, the Model Validation Peer Review Charge did not. EPA nonetheless collected additional information in response to Reviewer comments where practicable. After 8 years of model development, EPA believes that appropriate and adequate data collection activities have been conducted, and combined with changes to model parameterization following input from the Reviewers, model development has been sufficient to achieve the goal of the modeling study.

EPA responses to assertions regarding specific processes are provided in the subsections below and in other General Topics, and supporting analyses and discussion are provided in the FMD.

1.1.4 Responses to Specific Comments on General Topic 1

W. Frank Bohlen

The global list of processes and governing factors incorporated in the linked three (3) model system is comprehensive and includes all those necessary for a detailed evaluation of PCB transport, fate and bioaccumulation (see Table 2-1 pp.2-3 Vol.1). In addition, the model, as structured provides an excellent framework for the systematic evaluation of each of the factors governing PCB transport and its ultimate bioavailability in the Housatonic River. This framework directly complements quantitative study of transport and subsequent investigation and ranking of remedial alternatives.

Response 1-FB-1:

EPA agrees with the Reviewer’s comments that the model framework is comprehensive and includes all processes necessary to investigate potential remedial alternatives.

Despite the completeness of the global list, however, realization of the full potential of the models is governed by the extent to which model formulation provides accurate process simulation. Review indicates that model utility would benefit from improvements/modifications in a number of areas.

Response 1-FB-2:

EPA disagrees with the Reviewer's assertion that further study would necessarily result in a demonstrably superior model framework. See General Response above.

Moving to the areas of standing water in the main channel, the backwaters and Woods Pond, I remain concerned about the accuracy of the sediment transport formulations. This concern is not entirely alleviated by the sensitivity analyses presented in the previous calibration report (Weston, 2004) and those included in the MVR. The response to a 50% variation in a variety of parameters overall seems reasonable and makes clear that all of the model results are sensitive to upstream boundary conditions. This, of course, is not surprising given the role of streamflow across this boundary in model dynamics and the limited number of areas within the PSA sufficient to serve as a sediment supply. What's demonstrated is in fact why one worries about the accuracy of the sediment transport formulation.

Response 1-FB-3:

The model's ability to simulate floodplain interactions is well documented. The model reproduces the extent of flooding documented during extreme events (Hurricane Bertha). Further evaluations of model performance based on conservation of mass (continuity) demonstrate that the model correctly simulates the storage (and release) of water from flooded areas during out-of-bank events. The hydrodynamics of floodplain interactions controls the movement of sediment and PCBs between the river channel and floodplain. The theory for describing sediment transport and the formulations application are identical (and equally as accurate) regardless of whether the material in transport is in the channel or the floodplain. The mass flux ("process-based flux" and "tracer runs") summaries of the model results illustrate the connection between the channel and floodplain areas in the model. See FMD Appendix B.7 and FMD Section 4.2.

Nor are concerns alleviated by model runs requiring what appear to be inordinately high diffusive fluxes to explain the simulated increase in PCBs at New Lenox road relative to those at Holmes Road (see pp.6-72 and Fig.6.2-42 MVR) during low flow conditions. This response suggests that the calibration of the sediment transport formulation might, because the majority of the available data were obtained during high flow events, have produced a function that is overly specific to higher flow conditions. This calls to mind the comments of Dr. Lick regarding the need to reduce the number of adjustable parameters in the transport formulation and more accurately specify those that remain (comments that I second) so as to have confidence that the algorithm is an accurate representation of governing physics and not simply some curve fitting routine.

Response 1-FB-4:

At the outset, it is worth noting that this Reviewer opined that the PCB sediment-water mass transfer ("diffusive") flux was inordinately high, while, in contrast, other Reviewers commented that the mass transfer flux was "too low." Further discussion of the mass transfer flux is provided in the response to General Topic 1B and in FMD Appendix B.10.

The number of adjustable parameters in EFDC was limited to the fewest number possible. Most sediment transport parameters depend on the measurable physical properties of the sediment (grain size, density, etc.) and are not adjustable. In fact, only parameters for cohesive sediment transport were even potentially adjustable. Even in these few instances, the parameters were derived from site-specific measurements. The model results are a reflection of the proper representation of the physics of sediment (and PCB) transport and are not a simple curve fit of data. This is illustrated in Figure 3.3-1 of the FMD, and further discussion is presented in FMD Section 3.3.2, Section 4.3, and Appendix B.8.

I would recommend that increased focus be placed on the sediment transport formulation and that model runs be conducted in which the sediment supply across the upstream boundary is set to zero. Minimal "tuning" and reasonable results under these conditions would increase confidence in the formulation and directly benefit future evaluations in the CMS that very likely will be dealing with transport in specific areas within the PSA and require accurate estimates of local mass movements.

Response 1-FB-5:

EPA believes that it is unrealistic to assume that solids loadings at the upstream boundary would be zero under any circumstance, or that any model results generated using such an assumption could be considered reasonable. However, a sensitivity analysis was performed where the upstream solids loads were increased and decreased by 50% and the resulting model output was presented (see Section B.3.4 of the Model Calibration Report [MCR], Section 5 of the Model Validation Report [MVR], and Section 3.7 of the FMD). EPA believes that the Reviewer was inquiring if the spatial erosion and deposition patterns simulated by the model appear reasonable. Such analyses were presented in Figure 4.2-51 of the MVR for a range of flow conditions. These results were judged reasonable by the modeling team.

Next, additional work is required to develop an accurate formulation of sediment transport. The suggestions of Dr. Lick in terms of both equation parameters and the structure of the sediment bed should be carefully evaluated. There seems to be an abundance of data and experience to suggest that 6in is an overestimate of the active bed thickness. There is also concern that the model as it exists may be biased to high flow conditions. Add to these questions regarding side bank erosion and floodplain dynamics.

Response 1-FB-6:

EPA disagrees that further work is needed to develop accurate sediment transport formulations. For non-cohesive sediments, the transport formulas used are well described in the peer-reviewed literature. For cohesive sediment, the transport formulas used are those recommended by Dr. Lick. Further, cohesive sediment erosion parameters were derived from site-specific measurements made with the Sedflume. These data have been evaluated in accordance with suggestions made by Dr. Lick.

It should be noted that the active thickness used in the model was determined based on an extensive set of physical measurements, literature review, and site-specific sediment profile images. See FMD Appendices A.3, B.4, and B.5.

Douglas Endicott

There appear to be parameterization errors in specific sediment transport and PCB fate and transport processes.

Response 1-DE-1:

See General Response above. EPA believes that the processes in the models were correctly parameterized and all relevant processes accurately represented. This is clearly documented in the FMD.

Marcelo H. Garcia

While a substantial effort has gone into improving the capabilities of the Housatonic River Model since its calibration, it is clear that there is still a long way to go before the model can truly be used as a "predictive tool" to quantify future spatial and temporal distributions of PCBs (both dissolved and particulate forms) within the water column and the bed sediment.

This reviewer is of the opinion that the Housatonic River Model will need to be continuously improved until all the processes (biological and physical bed sediment mixing, streambank erosion, floodplain deposition, etc.) relevant to the transport and fate of PCBs are not only accounted for in the model but are also well represented and based on sound knowledge of sediment transport mechanics, stream biology, ecology and morphodynamics.

Response 1-MG-1:

EPA disagrees that the model cannot be used as a "predictive tool" and must undergo continuous improvement. Eight years of model development have resulted in a model framework, as calibrated and validated, that is adequate to achieve the modeling study goal to simulate future conditions and the relative performance of potential remedial alternatives. The Reviewer's comment is based on a false premise that the balance of fate processes is not correctly represented. See the General Response above.

Given the facts that several processes are yet to be fully characterized in the model and that a coarse computational grid has been used for the model calibration, I do not think that a full sensitivity analysis can be conducted at this time.

Response 1-MG-2:

EPA disagrees with the Reviewer's assertion that several processes have yet to be characterized. All relevant processes needed to simulate PCB fate, transport, and bioaccumulation have been adequately characterized. Further, the model sensitivity to parameter perturbation has been evaluated and was presented in the MCR as well as MVR Section 5. The sensitivities of the model parameters are further described in FMD Section 3.7.

Frank Gobas

The temporal response of the model is based (i) on the parameterization of the sediment-water exchange rate of PCBs (which is the rate determining step in the depuration of the River) and (ii) on model calibration. With regards to the parameterization of the sediment-water exchange of PCBs, there are several model parameters including the depth(s) of "accessible" and "non-accessible" bottom sediments, diffusion rates, bioturbation rates and subduction velocities that are currently difficult to measure or assess. The best strategy for parameterization is to use the best empirical data and best expert judgment possible. I am not convinced that this was achieved in this study. Dr. Lick has provided several papers to demonstrate that the current parameterization of the sediment-water exchange process in the model is not consistent with some key studies and observations. Also, I place considerable weight on Dr. Lick's expert judgment and his lack of confidence in the selection of parameter values. However, that said, it is unknown at this time whether an alternative parameterization of the sediment-water exchange processes will produce a significantly different and better characterization of the temporal response of the PCB concentration in the River.

Response 1-FG-1:

EPA agrees with the Reviewer's conclusion that an alternative parameterization of the processes may not result in a significantly different or better characterization of the temporal response of PCB concentrations in the Housatonic River. This is demonstrated by the extensive exploratory analyses conducted in response to the Reviewers' comments, many of which were discussed in prior documents and all of which are provided in the FMD and/or discussed in this Responsiveness Summary. Throughout the model development process, great care was exercised to ensure that each model process was parameterized in an independent manner to achieve the proper balance of fate processes. See FMD Section 3.3.2, Section 4.3, and supporting appendices.

The modelers should explore alternate parameterization schemes (possibly with the help of Dr. Lick) of the sediment-water exchange processes with the goal to select model parameterization schemes that are defensible based on available laboratory and field observations.

Response 1-FG-2:

Sediment-water mass transfer was parameterized from site-specific measurements. See FMD Appendix B.10.

With regards to achieving model objective #2, i.e. ability to quantify the historical and current relative contributions of various PCB sources to PCB concentrations in water and bed sediment, I do not think that the model reasonably accounts for the relevant fate processes. The reason is that the model has difficulty assessing the amount of historical PCB mass that is “accessible” by the River. Difficulties in the selection or determination of an “accessible” sediment layer and decisions regarding the inclusion or exclusion of flood plains as accessible sources of PCBs to the Housatonic River contribute to the overall difficulty in assessing the current mass of PCBs in the River. As a result it is difficult to assess the relative contribution of any current inputs of PCBs into the River. The lack of any significant change in the PCB concentrations in the River over time over the period that significant removal of PCB sources in the immediate vicinity of the GE facility remediation took place may be an indication that historical sources of PCBs throughout the PSA are likely the main contributor to current PCB concentrations in the River.

Response 1-FG-3:

As described throughout this General Response, the FMD, and all other supporting documents, EPA believes that the model reasonably accounts for all relevant fate processes, that each individual transport and fate process was correctly parameterized, and that the relative importance (balance) of each process is appropriately represented. The Reviewer’s comment regarding an “accessible” sediment layer was unclear and interpreted to mean the bioavailable sediment depth for estimating food chain exposures. The bioavailable depth was determined based on an extensive set of physical measurements and literature review, and was verified using site-specific sediment profile images (see FMD Appendices A.3, B.4, and B.5). The river floodplain is an important reservoir of PCBs in the Housatonic River. Risks to both human and ecological receptors were identified from exposure to PCBs in the floodplain. Therefore, EPA believes that no simulation of PCB transport and fate in this system is complete without explicit consideration of the floodplain and its interactions with the main river channel. As part of his final point, the Reviewer is correct that there has not been a significant change in sediment PCB concentrations within the PSA over time. However, in making this point, the Reviewer incorrectly assumed that the period over which no statistically significant change in PCB concentrations in sediment was observed corresponds to the timeframe during which sediment and bank soil remediation was performed upstream in the vicinity of the GE facility. Nonetheless, EPA agrees with the Reviewer’s conclusion that sources of PCBs within the PSA are likely a main contributor to PCB concentrations in the river, as shown in the tracer analyses (see Section 4.2.5 of the FMD).

The model is the only available tool to simulate the future response of PCB concentrations in response to remediation efforts. The model framework represents a suitable approach to estimating the future time response of PCB concentrations in the River and the calibration and validation of the model have involved significant efforts. The most valuable information for the calibration and validation of the model is a change in PCB concentrations in the River in response to a known reduction in PCB loading. This kind of information was not obtained in the current study as concentrations of PCBs in the River showed little or no significant variation with time. The current model therefore has to rely on the characterization of a number of key state variables for the estimation of the long term temporal response of PCB concentrations to remedial scenarios. The key parameters include the amount of "available" River and floodplains sediments and rates of resuspension, diffusion, bioturbation and subduction. All of these model state variables are either currently unmeasurable or very difficult to measure or estimate. As a result the model's outcome with regards to the long term time response of PCB concentrations in the River is uncertain. The model uncertainty translates in considerable uncertainty about future PCB concentrations in the River resulting from remediation efforts or the absence of remediation.

Response 1-FG-4:

EPA agrees with the Reviewer's comment that little change in PCB concentrations (no statistically significant trend) was observed during the period of data record, thus making the assessment of model performance more challenging and introducing uncertainty into the model simulations. However, even without changes in PCB concentrations, there is already uncertainty associated with model predictions, given that the data demonstrate orders of magnitude variability in PCB concentrations over the same period in space and time. EPA also agrees with the Reviewer that many of the variables in the model are either currently unmeasurable or very difficult to measure in a way that would result in a decrease in uncertainty. Where possible, EPA did undertake such measurements, and where not possible, EPA did extensive literature reviews to derive the best possible estimates for such values.

Wilbert Lick

1. In the present model, the processes of sediment erosion, sediment deposition, the finite sorption rate of PCBs by the sediments, and the sediment-water flux due to "diffusion" are described incorrectly and inaccurately. This is exacerbated by a very coarse numerical grid used to define the bathymetry of the river and an unnecessarily fine grid to define the bathymetry/topography of the floodplain. More specific reasons for these comments as well as specific suggestions for improvement of the model are presented in the complete response which is attached. I do not believe that the present model can reasonably account for the relevant processes affecting PCB fate, transport, and bioaccumulation in the Housatonic River to a degree consistent with achieving the goal of the modeling study.

Response 1-WL-1:

See General Response above. In this comment, the Reviewer made a series of flawed assumptions that resulted in erroneous conclusions regarding the model representation of processes. The majority of processes in the model were parameterized from site-specific measurements.

> For example, the model representation of PCB partitioning is entirely consistent with site data. However, the Reviewer asserted that it was necessary to include time-dependent sorption in the model because use of the equilibrium partitioning approach would necessarily result in an improper representation of PCB phase distributions, particularly during high-flow events. EPA evaluated the Reviewer's proposed methods, conducted simulations, and documented that: (1) use of equilibrium partitioning yields PCB concentrations that are within 5 to 10% of results obtained using time-dependent ("frozen reaction rate") sorption; and (2) PCB fluxes computed used equilibrium partitioning during an event are nearly identical to those computed using "frozen reaction rates." This small difference in model results is less than the uncertainty associated with reported PCB sorption rates. These analyses are further described in FMD Appendix B.9.
>
> Responses to assertions that other processes were incorrectly parameterized are provided in the remaining sections of this Responsiveness Summary. Complete documentation of the parameterization of these processes is provided in the FMD and appendices.

In the long term, the main source of contaminants in the Housatonic River is the bottom sediments (both those in the river and in the floodplain). In this long term, environmental conditions will change with time and will be different than they are at present, especially because of and after remediation. For purposes of predicting water quality, it is therefore essential to accurately determine not only the flux of contaminants between the bottom sediments and the overlying water but also the parameters on which this flux depends. Otherwise, long-term predictions of water quality will not be accurate or believable. Because of this, my comments will emphasize processes which govern the sediment-water flux of contaminants, i.e., sediment erosion, sediment deposition, the sediment-water flux due to "diffusion", and equilibrium partitioning.

Response 1-WL-2:

> EPA agrees. Model simulations indicate that river sediment is likely to be a major source of PCBs.

In my last review (Model Calibration, May 13, 2005), I commented extensively on the proposed calibration of both sediment erosion and deposition by means of the measured suspended solids concentration, C. A simple example was given whereby it was easy to see that a numerical model can "predict" the observed values for C with an almost arbitrary value of erosion rate as long as the deposition rate was changed accordingly, i.e., such that the two were equal and gave the observed C. I stated "For a predictive model, the values of erosion rate and deposition rate can not both be determined from calibration of the model by use of the suspended solids concentration alone." I later stated 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 lead to the incorrect predictions of long-term behavior."

EPA responded by more-or-less agreeing with the above statements but then stating “This concern does not recognize that there is a constraint imposed by PCB transport that results from resuspension and deposition processes.” Although this argument has some validity, it is not sufficient or correct as I will argue below. This problem of non-unique solutions is important, is more general than that indicated above, and is related to the necessity for accurately determining the basic processes that govern the sediment-water flux of PCBs and other hydrophobic organic chemicals (HOCs). Because of this, I will return to this problem of non-unique solutions after discussing the flux processes mentioned above. Accurate descriptions of the basic processes also depend on an adequate resolution of the bathymetry/topography of the Housatonic. Because of this, comments and suggestions on the problem of numerical gridding in the model will be made. Some discussion on unexplained results of the present model will then be given. A summary and specific suggestions for improvements to the model will conclude my comments.

Response 1-WL-3:

EPA disagrees with the Reviewer’s comment that the model parameterizations are “unconstrained.” The calibration of sediment transport included consideration of suspended solids concentrations as well as the net deposition flux to the sediment bed. The net deposition flux was estimated from cesium-137 data as described in MVR. Although more qualitative than the analysis of cesium-137 data, computed bed shear stresses provide another “constraint” on cohesive solids deposition because solids will not be incorporated into the sediment bed if the shear stress exceeds the critical shear for deposition. As the Reviewer acknowledges, the final constraint on the parameterization of erosion and deposition is successful simulation of PCB concentrations. This includes PCB concentrations in both the water column and the sediment bed. The model parameterization is appropriate because suspended solids, net deposition fluxes, water column PCBs, and the rate of change of PCBs in the sediment bed are all correctly simulated over time as demonstrated by the evaluation of model results presented in FMD Section 3.5 and Section 4.

However, the Reviewer’s comments regarding the “uniqueness” of model solutions are unrealistic. These comments are based on the assumption that a so-called “unique” solution must be the goal of a modeling study and that without a “unique” solution the results obtained will necessarily be incorrect or inaccurate. In an “algebraic” sense, a unique solution can only occur when there is a unique set on inputs such that there is a one-to-one correspondence between model inputs and outputs. For practical applications of environmental models, the Reviewer’s position is unrealistic because naturally variability in model inputs such as upstream solids and PCB loads, sediment bed PCB concentrations, and partition coefficients preclude development of a unique set of inputs.

In this light it is more realistic to consider whether model results are sufficiently definitive to provide a meaningful solution. In this context, model results can be considered “sufficiently definitive” if they describe the essential characteristics of the natural system with an accuracy and resolution that is adequate to achieve

> the goals of the modeling study. Pre-determined (a priori) model performance measures were used to assess if model results provide a sufficiently definitive description of the Housatonic River. As documented in FMD Section 3.5 and elsewhere in the FMD, model results are, in virtually all cases, well within the established performance measures. For this reason, EPA believes the model framework is adequate to achieve the goal of the modeling study to simulate future conditions and the relative performance of potential remedial alternatives.

As discussed above, the processes which govern the sediment-water flux of HOCs (sediment erosion, sediment deposition, the sediment-water flux due to "diffusion", and equilibrium partitioning) are described incorrectly and inaccurately. Each of these processes can modify the flux by factors of two to ten. Nevertheless, EPA documents indicate that there is good agreement between the calculated and measured suspended solids concentrations as well as contaminant concentrations. At the same time, sensitivity and uncertainty analyses also seem to say that the model is doing a good job. **How can this be? Are accurate models of sediment and contaminant transport and fate really unnecessary?**

The answers to these questions are in the non-uniqueness of calibrated solutions and the nature of sensitivity and uncertainty analyses. More specifically, it seems that a modeler can assume a wide range of parameters to describe a particular process and still "calibrate" the model so as to determine a mathematical solution which agrees with observations over some time interval. As discussed above, examples of parameters which significantly affect flux processes are (1) different values of the power n in erosion formulas (this gives greatly different erosion rates at high shear stresses depending on the value of n), (2) different parameterizations for settling speeds, (3) different process models and parameters for the sediment-water flux due to "diffusion", and (4) equilibrium partitioning (equivalent to high reaction rates), frozen reaction rates, or anything in between. Calibration of a model does not guarantee that the processes in the model are described properly. At the risk of being repetitive, a water quality modeler can always get good agreement between calculated and observed quantities for a limited time interval and limited conditions, whether the fundamental processes are described properly or not. Another modeler, with quite different descriptions of processes and/or different parameters in his/her model, can get equally good agreement between the calculated and observed quantities. However, future predictions by the different models and modelers will be quite different. This has been demonstrated repeatedly (e.g., see comments on Fox River modeling (Tracy and Keane 2000) in my comments of May 2005). In other words, **calibration is necessary but not sufficient.**

Response 1-WL-4:

> See Response 1-WL-3 and responses to the remainder of General Topics 1A through 1D.

E. John List

My opinion is that the underlying processes involved with the PCB fate and transport do not appear well enough understood for a meaningful model to be developed. If the processes that seem to define the PCB fate and transport are not included in the model then their sensitivity cannot be assessed.

> **Response 1-JL-1:**
>
> EPA believes that the underlying processes controlling PCB fate and transport have been adequately characterized to construct a credible model framework. These processes are also understood in sufficient detail at the spatial and temporal scales of interest for the intended use of the model. While mechanistic representation of factors controlling the variability of PCB concentrations or other constituents across small spatial scales may be of interest from a research perspective, the model framework does not need to explicitly represent such small variability to achieve the goal of the modeling study (see Sections 2, 3, and 4 of the FMD).

Since the basic processes for the fate and transport of PCB are clearly not properly included in the model their uncertainty cannot be assessed.

> **Response 1-JL-2:**
>
> See General Response above.

1.2 GENERAL TOPIC 1A: EROSION AND DEPOSITION

1.2.1 Comment Summary

Reviewer Name	Page Number/Line Number
Adams	Page 5, lines 1 to 12
Bohlen	Page 10, lines 3 to 14
	Page 10, lines 33 to 34
	Page 12, lines 6 to 13
Endicott	Page 3, line 28 to page 4, line 3
Lick	Page 4, line 12 to page 5, line 43
	Page 6, lines 1 to 7
	Page 12, lines 21 to 26
	Page 12, lines 28 to 32
List	Page 1, lines 34 to 39

1.2.2 General Topic Summary

Erosion of non-cohesive and cohesive sediment occurs through different mechanisms. Non-cohesive particles behave independently and act as individual particles. Cohesive particles interact with other particles and act as aggregates. Different formulations for these processes are used in EFDC. Garcia and Parker (1991) and Wu et al. (2000) are used to describe non-cohesive sediment erosion. Lick et al. (1994) is used to describe cohesive sediment erosion.

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 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.

Many of the Reviewers concurred with the opinion expressed by a particular Reviewer that *n* (the cohesive sediment erosion exponent) should be greater than the value used in the modeling study, most agreed with the particular Reviewer that it should equal 2 or more.

Two Reviewers commented on the use of the deposition rate as a calibration parameter, and one Reviewer recommended the use of his flocculation theory for deposition.

1.2.3 General Topic Response

For the EFDC application to the Housatonic River, the erosion of non-cohesive sediments is represented by relationships for suspended load and bedload as described by Garcia and Parker (1991) and Wu et al. (2000). These relationships were chosen because they are applicable to sediment mixtures, include terms to describe the hiding and exposure of particles within the bed, and have been validated by comparisons to experimental measurements and field data.

In contrast, the erosion of cohesive sediment is highly variable and often requires site-specific information regarding the characteristics of the bed. For the EFDC application to the Housatonic River, the parameterization of cohesive sediment erosion was based on site-specific data.

Experiments were conducted to characterize the erosion properties of the sediment of the Housatonic River (Gailani et al., 2000). Two types of experimental devices were used. The Sedflume (McNeil et al., 1996; Jepsen et al., 1997) was used to measure erosion rates as a function of shear stress and depth in the sediment. Gailani et al. (2000) presents a detailed description of the experimental procedures and data.

The Sedflume experiments quantify erosion rates at imposed shear stresses, but the data also provide a description of how the shear stress required for erosion changes with depth in the sediment bed. The absence or occurrence of erosion through a range of imposed shear stresses can be summarized by two values: the maximum shear stress imposed without initiating erosion, and the minimum imposed shear stress that caused erosion. The shear stress required to initiate erosion, referred to as the critical shear stress for erosion, lies between these two values.

In response to Peer Review comments, the cohesive sediment erosion exponent (n) for the EFDC application to the Housatonic River was set to a value of 2.

A complete discussion of the analyses and evaluation of the critical sheer stress for erosion is provided in Appendix B.8 to the FMD.

1.2.4 Responses to Specific Comments on General Topic 1A

Eric Adams

The erosion formulation and parameters come from analysis of SedFlume data. W. Lick argues that the exponent n (denoting dependency of erosion on shear stress) should be significantly greater than the chosen value, which would produce proportionally more erosion during high flow conditions. I agree with him, but am also leery of the fact that shear stresses in the model are computed based on a grid width that is essentially equal to the river width. Thus the computed hydrodynamics will yield cross-sectional average velocities, which ignores regions within a cross-section with relatively high velocity and erosion (refer to later discussion under Question 7). Unlike conditions in SedFlow, the real river is non-uniform. Thus the formulation of erosion can not be divorced from the question of grid resolution. Indeed, it would seem incorrect to simply import an erosion calculation from SedFlume.

Response 1A-EA-1:

As described in FMD Appendix B.8, the parameterization of cohesive sediment erosion in EFDC was developed from the site-specific Sedflume measurements for the river. As demonstrated in Appendix B.8, the model results are not particularly sensitive to the erosion exponent value for the range of shear stresses encountered. For the final model simulations, EPA used a value of $n = 2$ for cohesive sediment erosion as recommended by the Reviewers.

As noted by the Reviewer, there can be a connection between model grid scale and the parameterization of model processes. In the case of cohesive sediment erosion, the variability of Sedflume measurements from individual core locations was large. Given this variability, the most reliable approach to using Sedflume measurements in the model framework was to compute erosion parameters on a reach-wide average basis. This approach is appropriate in consideration of the intended spatial scale of use for the model framework.

Also as noted by the Reviewer, it is difficult to simply import erosion parameters from Sedflume measurements without consideration of grid scale. To account for possible grid scale dependencies, the critical shear stress for cohesive sediment erosion in EFDC was calibrated within the range of the Sedflume measurements. Further, model sensitivity to the erosion rate coefficient was evaluated as presented in the MVR. This sensitivity analysis was subsequently expanded to include both the erosion exponent and the erosion rate coefficient as documented in FMD Section 3.

W. Frank Bohlen

The TSS data set appears to be moderately robust although I worry that too much emphasis might have been placed on storm conditions. The data also appear to be primarily point measurements with some representing integrated measurements over the vertical. The absence of time series data prevents detailing of processes such as the time scales of resuspension and deposition during rising or falling stage with particular emphasis on the onset of resuspension during relatively low flow conditions. The influence of such low level but persistent resuspension and transport on PCB fluxes is largely ignored in the present study which places primary emphasis on storm events. Absent the low flow details it's difficult to assess the role of these processes in the long term. Such assessments will be a subject of study in the upcoming CMS. As noted above I'd recommend immediate initiation of a monitoring program designed to provide time series observations of TSS concentrations at a number of selected stations throughout the PSA.

Response 1A-FB-1:

Time series data for TSS (including data collected on both the rising and falling limb of the hydrograph) were collected during the storm events, and when possible, data were also collected at the time of low flow preceding the rise of the hydrograph. EPA disagrees that the influence of low level but persistent resuspension and transport on PCB fluxes was largely ignored and that the primary emphasis was placed on the storm events. Both storm flow and low flow conditions were evaluated during the calibration process, and both must be considered in modeling potential remedial alternatives in the CMS. It should be noted that GE is expected to continue monthly monitoring for PCBs and TSS at many locations in the study area.

An additional check on model results that should be considered includes the use of time series bathymetric data to check on the accuracy of sediment erosion/deposition estimates.

Response 1A-FB-2:

The time series bathymetric data were evaluated for use in evaluating model performance. The obstacle preventing this type of analysis was that bathymetric data do not exist at enough locations along the stream and at an appropriate temporal frequency to permit accurate estimation of erosion and deposition volumes in a manner that could be appropriate compared to model results to provide a quantitative measure of model performance.

As noted on several occasions above, visual examination of the figures suggests that the model overpredicts TSS concentrations at low flows and underpredicts them at high flows. The distribution of residuals provides clear indication of the overprediction at low flows for most stations (see Table 6.2-6). Although a number of reasons for these differences (both high and low flow) are presented there is no mention of the possibility that they are the result of less than accurate formulation of the sediment transport process, in particular bed erosion/deposition, in the model. As discussed above, this seems more than possible and should be carefully evaluated since it will be of increasing concern during the CMS.

Response 1A-FB-3:

Sediment erosion in the model framework considers both non-cohesive and cohesive sediment. In EFDC, the erosion of non-cohesive sediment is computed using peer-reviewed transport formulations by Garcia and Parker (1991) and Wu et al. (2000). Cohesive sediment erosion in EFDC is computed using the formulation recommended by Dr. Lick (and other Reviewers) and is parameterized with site-specific data. While there is uncertainty associated with their predictive abilities, these formulations represent the best available tools to drive erosion calculations in the model. Further, it is worth noting that sensitivity analyses were performed for the parameters used in these sediment transport formulations.

Douglas Endicott

Sediment transport parameterization: Cohesive sediment resuspension is parameterized based on the analysis of data from SEDFLUME experiments conducted on a number of sediment cores. Results of this analysis produced shear stress exponents of 1.59 for sediments in Reaches 5A and 5B and 0.95 in Reaches 5C and 6[1.] These values are considerably lower than shear stress exponents reported in the literature. For example, Lick, Ziegler and coworkers have reported that for cohesive sediments tested at a significant number of sites, the shear stress exponent is generally constrained within a fairly narrow range (n=2.6±0.3). Based upon this, it appears that parameterization of cohesive sediment resuspension in the model is erroneous, and should be corrected consistent with the guidance offered by Dr. Lick. As a consequence, it will also be necessary to recalibrate deposition rates.

Response 1A-DE-1:

In response to Reviewer comments, cohesive sediment erosion in EFDC has been changed to a value of $n = 2$. However, it is worth noting that the erosion exponent values presented by the Reviewer ($n = 2.6 \pm 0.3$) were derived from the shaker (Particle Entrainment Simulator, PES) and not the Sedflume. Erosion exponents derived from the PES are not believed to be applicable to the Sedflume as has been noted by Dr. Lick. Furthermore, as documented in FMD Section 3, increasing the cohesive sediment erosion exponent in the model did not require recalibration of deposition rates This assessment was based on comparisons of water column TSS values and comparisons to net sedimentation measurements in Woods Pond (see FMD Appendix B.6).

---

[1] Model Calibration Report, Volume 3 - Appendix B Hydrodynamic and Sediment/PCB Fate and Transport Model Calibration, Attachment B.5 Analysis of Sediment Erosion Data

Wilbert Lick

In the previous review on Model Calibration (May 13, 2005), I commented extensively on sediment erosion. Although those comments are still valid, I won't repeat all of them here. However, I would like to repeat the following from those comments.

"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 $d \rightarrow 0$ 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."

One 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, sequential 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. I suggested an interpolation procedure that we had used before and which gave us reasonable results. EPA did not seem to have good results with this procedure. Attached is a description of a modified procedure which I have applied to several randomly selected cores on each of the Kalamazoo, Housatonic, and Passaic Rivers. This procedure is more fundamental and correct. In all cases, it produces an n that is equal to two or greater in Eq. (5) above, just as has been demonstrated by all Sedflume laboratory measurements that we have made. With this procedure, the coefficient modifying $(\tau - \tau_{cn})^n$ and n are functions of depth for each representative area core and are obtained from the Sedflume data. The fact that n is two or greater is important in determining erosion rates at high shear stresses, i.e., during big events, and can lead to shear stresses higher by an order of magnitude than the n's chosen by EPA for the Housatonic.

Response 1A-WL-1:

As understood by EPA, the central topic of this comment is that the cohesive sediment erosion exponent *n* should have a value of 2 (or more). In response to Reviewer recommendations, cohesive sediment erosion in EFDC is now parameterized using $n = 2$. See the General Response above. For further detail regarding the formulations and parameterizations for erosion that are used in EFDC, see FMD Section 2 and Appendix B.8.

In that previous report, I also stated the following. "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)."

EPA stated that the surficial layer was assumed to be 7 cm thick, not 6 inches as I stated. Since bed coarsening occurs in a layer only a few particle diameters thick, the assumption of a 7 cm mixed layer is also incorrect. **The above comments are still valid.**

Response 1A-WL-2:

For clarification, it should be noted that surface sediment layer thicknesses in EFDC range from 4 to 7 cm.

Bed armoring and the development of armor layers is a complex topic. Armoring depends on many factors including the grain size distribution of particles in the bed at the sediment surface. In EFDC, some of the effects of armoring are addressed through the hiding and exposure factors for the erosion of individual sediment grains in a non-uniform sediment mixture. In addition, the fraction of cohesive sediment in the sediment bed is also used to modify non-cohesive erosion, as described in FMD Section 2.3.2.

As noted by the Reviewer, an armor layer in the bed can form in a thickness of only a few (2 to 3) grain diameters. Given typical grain size distributions in the Housatonic River, the potential exists for armor layers to form over a depth of 1 to 1.5 cm, based on the average diameter of the NC3 solids class (d = 0.516 cm). However, this represents a minimum. A further limitation on the development on an effective armor layer is the extent and frequency of bed disturbance by turbulent shear and burst or other factors that contribute to sediment mixing. Bed load transport processes can also impact armor development.

Sediment profile images and bedform observations for the Housatonic River suggest that frequent sediment disturbances likely prevent the development of significant or persistent armor within the first few centimeters of the bed surface (see photographs in FMD Appendix A.3 and Appendix B.5). Based on these factors, EPA believes that the model representation of bed armoring is adequate to meet the goal of the modeling study.

I suggested the use of a dynamic flocculation theory that was recently developed and was relatively simple. EPA seems to have had problems implementing the theory. I presume from what they said that this was due to numerical stability problems. That’s too bad because it would have given reviewers more confidence in the modeling of deposition rates.

Response 1A-WL-3:

In response to Peer Review comments, a mechanistic flocculation model was tested for use in EFDC. Specifically, the approach described by Lick et al. (2005) was tested. This flocculation model includes a non-linear iterative solver for floc diameter, which controls the effective settling velocity. While tests conducted for steady flow and steady concentration conditions without erosion appeared promising, the approach of Lick et al. (2005) could not be used to simulate flocculation for normal river conditions. Normal river conditions are unsteady because flow and concentration can change rapidly over short periods of time. When tested for the unsteady conditions routinely expected to occur in the Housatonic River, when there is erosion and rapidly changing flow and suspended solids concentrations, the flocculation approach described by Lick et al. (2005) failed because it computed unrealistically large settling velocities that caused all cohesive sediment to settle almost instantaneously, causing subsequent computations of zero floc diameters. Such failure under unsteady conditions is not entirely unexpected because it is more difficult to obtain solutions to many non-linear equations when conditions change rapidly. Further description and details of the evaluation of the Lick et al. (2005) flocculation model are presented in Section 4.1.2 of the MVR.

> As a clarification, it is worth noting that the failure of the flocculation model was not the result of numerical stability (or instability) in EFDC.

In the present model, erosion rates are too low (n is too small). During big events, erosion rates may be as large as, or greater than, 10 times that predicted at present. The major effect of this will be on the maximum depth of erosion during big events. A better analysis of Sedflume results as suggested in the attached report will determine a higher and more reasonable value for n. The coefficients in Eqs. (3), (4), and (5) should vary throughout the river as required by the data.

> **Response 1A-WL-4:**
>
> See General Response above.

In the present model, deposition rates are also too low. Deposition rates are essentially a calibrated parameter and are determined such that the suspended solids concentration is calculated properly. If erosion rates are increased, deposition rates must also be increased in order to maintain good agreement between calculated and observed suspended solids concentrations.

> **Response 1A-WL-5:**
>
> As documented in FMD Section 3, increasing the cohesive sediment erosion exponent in the model did not require recalibration of deposition rates. This assessment was based on comparisons of water column TSS concentrations and comparisons to net sedimentation measurements in Woods Pond. See Response 1A-DE-1 and FMD Appendix B.6. It should also be noted that PCB concentrations in the sediment bed and water column provide an additional check on the appropriateness of gross erosion and gross deposition rates in the model. See General Response above.

E. John List

On the other hand, models to simulate the movement of sediment in a fluvial stream are significantly more problematic. The basic difficulty lies in the erodible stream bed. Bed erosion is uniform neither transversally nor longitudinally in the stream. In addition, the bed erosion and the consequent bed form changes result in a change in the stream friction factor, which feeds back into a modification of the hydrodynamics of the flow.

> **Response 1A-JL-1:**
>
> The Reviewer correctly describes the feedback between hydrodynamics and sediment transport. See General Response above.

1.3 GENERAL TOPIC 1B: SEDIMENT-WATER MASS TRANSFER

1.3.1 Comment Summary

Reviewer Name	Page Number/Line Number
Adams	Page 3, line 25 to page 4, line 43
Endicott	Page 4, lines 4 to 10
Garcia	Page 5, lines 27 to 31
	Page 6, lines 4 to 7
Gobas	Page 15, lines 29 to 38
Lick	Page 12, line 44 to page 13, line 14

1.3.2 General Topic Summary

In a general sense, the sediment-water mass transfer of contaminants is often described as a (non-hydrodynamic) diffusive flux and can be the primary mechanism for contaminant exchange between sediment and surface water during low-flow conditions when sediment from the bed is not expected to erode.

This sediment-water mass transfer occurs via the exchange of dissolved and bound phase PCBs in the sediment bed pore water with PCBs in the water column. This occurs by a number of processes, including pore water advection (i.e., groundwater exfiltration), diffusion, and bioturbation. All of these sediment-water exchange processes can be summed and collectively expressed in the form of a gradient-driven mass transfer.

Some Reviewers commented that the mass transfer coefficient was too small (or low), while others felt it was too large (or high).

1.3.3 General Topic Response

For the Housatonic River, PCB sediment-water mass transfer was determined from site-specific data. Specifically, mass transfer coefficients were determined from the spatial profile of PCB concentrations in the water column when flows are low and no sediment erosion is expected. For conditions when the diffusive flux of PCBs from the sediment bed is most important, river flows would be low to moderate and TSS concentrations would be expected to remain steady or decline while PCB concentrations would appreciably increase. Under these conditions, the mass transfer of PCBs from the sediment bed is expected to be the predominant fate process between Holmes Road and New Lenox Road because: (1) mass erosion of particles from the sediment bed is negligible (shear stresses are less than critical shear stresses for erosion); and (2) PCB loss from the water column via deposition is small because solids concentrations are relatively consistent. No specific temporal trend in mass transfer coefficients was identified. See the Responsiveness Summary to the Peer Review of Model Calibration and Appendix B.10 to the FMD.

Considerable effort was also devoted to analyzing the spatial patterns of TSS and PCB transport throughout the PSA. Longitudinal concentration profiles for composite and individual surveys at low, moderate, and high flow were assessed. Although spatial gradients in TSS concentrations are typically small, PCB concentration gradients measured between the confluence and New Lenox Road are large and identifiable. In contrast, TSS and PCB gradients farther downstream are smaller and more variable. Relative to other reaches, spatial patterns of TSS and PCB concentrations across Woods Pond are even more variable, reflecting the differences in the physical setting and corresponding hydrodynamics of the backwater effect of Woods Pond Dam. The assessment of spatial patterns was based on composite and individual concentration profiles for a wide range of river flows. See FMD Section 2.3.3 and Appendix B.10.

Based on consideration of measurements for 41 individual spatial surveys, geometric mean and mean values for the sediment mass transfer coefficient range from 1.04 to 1.72 cm/day. The calibrated mass transfer coefficient (K_f) value in EFDC is 1.5 cm/day and well within the range of values determined strictly from measured spatial gradients of TSS and PCB concentrations.

1.3.4 Response to Specific Comments on General Topic 1B

Eric Adams

The calibrated mass exchange coefficient is K_f = 1.5 cm/d, which actually seems to be on the small side, since it incorporates a number of processes in addition to strictly pore-water diffusion. The coefficient for pore-water diffusion by itself should reflect the rate at which bioturbation brings PCB-sorbed sediment to the interface, the rate at which PCBs are desorbed to the porewaters, and the rate at which diffusion transports the dissolved PCBs into the overlying water column. These processes work like resistors in series. Chen (1993) showed that the sediment-water flux can be expressed as

$$J = \frac{C_{sL} / K_p}{\dfrac{\delta}{D_m} + \dfrac{2.2R}{(1-\phi)[(D_b + D')D_m \rho_s K_p]^{1/2}} + \dfrac{L}{(1-\phi)\rho_s K_p D_b}} \qquad (6)$$

where C_{sL} is the sorbed phase PCB concentration at the depth of the mixed layer L, δ is the water side boundary layer thickness, and D' is the effective diffusivity of PCBs within the sediment (molecular diffusion as affected by tortuosity).

Jorgensen and des Marais (1990) suggest $\delta = 0.02$ to 0.1 cm, and Shaw and Hanratty (1977) suggest

$$\delta = \frac{12 D_m^{\ 0.3} \nu^{0.7}}{u^*} \qquad (7)$$

where ν is the kinematic viscosity of water ($\sim 10^{-2}$ cm^2/s) and u* is the friction velocity. Using $u^* = 0.25$ cm/s gives $\delta \sim 0.05$ cm. Assuming L = 5 cm, $K_p = 10^5$ cm^3/g, $D_m = 0.5\text{x}10^{-5}$ cm^2/s, $D' = 0.3\text{x}10^{-5}$ cm^2/s, $D_b = 5\text{x}10^{-7}$ cm^2/s (EPA's surface value), $\rho_s = 2.5$ g/cm^3, $\phi = 0.6$, and R = 0.01 cm, gives values for the three terms in the denominator of Eqn 6 of roughly 10^4, $3\text{x}10^1$ and 10^2 s/cm respectively. This suggests that the flux is indeed water-side controlled (i.e, biomixing supplies contaminant to the interface sufficiently fast, and the contaminant desorbs sufficiently fast, that diffusion on the water side limits the transport) and that the last two terms can be ignored. Even if D_b were smaller by an order of magnitude ($D_b = 5\text{x}10^{-8}$ cm^2/s), the three terms in the denominator would be 10^4, $3\text{x}10^1$ and 10^3 s/cm, leading to similar conclusions, albeit by a smaller margin. The second term represents the "resistance" due to desorption, and the fact that it is small suggests that equilibrium partitioning can indeed be assumed in computing the flux. The reciprocal of δ/D_m is K_f which, with the above numbers, is $\sim 10^{-4}$ cm/s (8.6 cm/d). This is in the range of the values computed directly from EPA's flux analysis (e.g., Figure B.4-30 of the MCR, which includes values between 0.8 and 250 cm/d, with the majority between 3 and 10 cm/d). However, it is significantly above the value of 1.5 cm/d identified in the Phase 1 model calibration. One possible reason for the calibrated value being significantly lower is that the PCB concentrations used for the upstream model boundary conditions are generally higher than the data, at low flow, which may cause the calibration to underestimate the sediment-water exchange flux downstream.

Response 1B-EA-1:

EPA disagrees with many aspects of the Reviewer's analysis, particularly the comment that the "low" mass transfer coefficient in the EFDC is the result of an error in assigning the model upstream boundary conditions. See General

Response above, FMD Appendix B.10, and the General Response for the balance of fate processes in General Topic 1.

As noted by the Reviewer, the sediment-water mass transfer process is used to represent the composite effect of individual processes that cannot be readily measured or otherwise quantified in a practical manner on an individual basis. As a practical consideration, the value of the sediment-water mass transfer coefficient (K_f) used in EFDC was determined by calibration. However, this calibrated model value is well within the range of values that can be determined strictly from measured spatial gradients of TSS and PCB concentrations using site-specific data for the Housatonic River from 41 surveys.

Based on these site-specific measurements alone, average K_f values for the Housatonic River range from 1.04 to 1.72 cm/day. The measured average K_f values do not depend on any simulation results or on the upstream boundary condition assigned in the model. Consistent with model development and evaluation procedures, EPA has relied extensively on site-specific data. These data all demonstrate that the model framework accurately describes the sediment-water mass transfer process for PCBs. On this basis, EPA believes the Reviewer's comments that the calibrated K_f value in the model is low and is impacted by upstream boundary conditions are incorrect.

Although the flux analysis indicated significant temporal variability in K_f that seems like it is correlated with stream flow rate, EPA's values of K_f based on complete model calibration appeared to be independent of flow rate. There seemingly should be some be some dependence, since increased flow would increase stream turbulence, decreasing δ and increasing K_f. As EPA acknowledges, as flow rate increases, it dilutes the water column concentration of PCBs, making it difficult to test for flow dependent fluxes. But observations do show that the model over-predicts water column PCB concentrations during low flow (when diffusive fluxes would dominate) which could reflect, at least in part, the lack of flow-dependence.

Response 1B-EA-2:

The extent to which the mass transfer coefficient depends on flow was explored through regression analyses. As a function of flow, the diffusive flux could range from 0.4 to 2.5 cm/day over the range of flow from 15 to 500 cfs, based on the spatial surveys between Holmes Road and Woods Pond Headwaters. However, it should be noted that the regression with flow explained only 22% of the variability in the measured K_f values (i.e., $r^2 = 0.22$) and that the range of K_f values (0.4 to 2.5 cm/day) is within one standard deviation of the average values determined from the spatial surveys. This range of K_f values is expected to be within the measurement error of the spatial survey data. Given these factors, EPA believes that the mass transfer coefficient in EFDC is properly defined without further representation of potential flow dependencies.

In addition, it must be noted that the Reviewer's comment regarding over-prediction of water column PCB concentrations due to the potential lack of flow dependency is not correct. The difference between simulated and measured PCB concentrations at low flow is the result of the variability of the TSS data on which the model upstream boundary condition, including the PCB boundary condition, is based. See FMD Appendix B.2 and General Topic 3.

The time scale for natural recovery is the mass inventory per unit area, $C_s(1-\phi)\rho_s L$, divided by the flux, $J = C_{sL}K_f/K_p$, or $\tau = (1-\phi)\rho_s L K_p/K_f$. For $K_f = 10^{-4}$ cm/s (my value), and using other parameter values from above, the time scale is 160 years. Thus diffusive exchange may not be very important for contaminated sediments that are buried at or below the assumed level of bio-mixing (i.e., most existing sediments). However, for recently transported sediments with smaller L (e.g., following remediation), the time scale could be much smaller, especially if K_f were even larger. For example, if L were only 1 and K_f were simply twice the above value ($2x10^{-4}$ cm/s), the time scale would be reduced to 16 years. However, the model would not be able to resolve the resulting sharp gradients with the current, relatively coarse, vertical grid scheme.

Response 1B-EA-3:

The concept of the time scale for natural recovery by PCB sediment-water mass transfer described in this Reviewer's comment is important. However, EPA disagrees with the Reviewer's conclusion. It should be noted that the K_f value used by the Reviewer was 10^{-4} cm/s (8.64 cm/day). This value is 5.76 times greater than the K_f value used in EFDC (based upon site-specific data). Consequently, when K_f = 1.5 cm/day is used in the Reviewer's calculations, the estimated response of 160 years becomes 900 years. Using a K_f value that is twice as large as the value currently in EFDC (i.e., 3 cm/day rather than 1.5 cm/day) would still result in a response time of 90 years, rather than 16 years as projected by the Reviewer. Under these conditions, the model grid has the resolution needed to represent any gradients in the sediment bed that could reasonably develop over time.

Douglas Endicott

The diffusive flux (i.e. sediment-water exchange) of PCBs from sediment pore water to the overlying water column was parameterized as a constant mass transfer coefficient during Phase 1 calibration. Phase 2 calibration has altered the spatial profiles of PCB water column concentrations at low flow conditions, suggesting that sediment-water exchange should be recalibrated. In all likelihood, the mass transfer coefficient should be increased, and made a reach-specific parameter via a relationship between the mass transfer coefficient and the benthos density.

Response 1B-DE-1:

See General Response above.

Marcelo H. Garcia

Model predictions in Figures 6.2-45 and 6.2-46, show that during low flow conditions the model does not capture observed longitudinal gradients in the water column. This results in the model underpredicting PCB fluxes from the sediments along Reach 5A, which in turn could lead to an under estimate of the bed sediments PCB contributions to fish in the Food Chain Model.

Response 1B-MG-1:

EPA disagrees with the Reviewer's comment regarding the ability of the model to simulate spatial trends. MVR Figures 6.2-45 and 6.2-46 show that the model appropriately represents the observed spatial gradients in the water column for low-flow conditions. To properly interpret these figures, it must be realized that the model performance envelope (shown in gray) and the data range (shown as the bars for the measurements) overlap. The overlap of the model and data ranges indicates that model performance is adequate.

It is important to recognize that the balance of fate processes in the model is properly represented (see FMD Section 3.3 and the General Response to General Topic 1). Further, the bounding analysis demonstrates that FCM results are not impacted by potential bias in PCB exposure concentrations at low flow. See FMD Section 4.3.3 for bounding analysis results and discussion.

In the case of low flow conditions, the model predictions seem to be very sensitive to diffusion parameters, suggesting the representation of pore water diffusion as well as the thickness of the so-called "mixed layer' need to be carefully analyzed.

Response 1B-MG-2:

EPA agrees with the Reviewer that the mass transfer flux needs to be carefully analyzed. The mass transfer coefficient used in EFDC is consistent with site-specific data. See the General Response above.

Frank Gobas

The importance of a diffusive flux of PCBs from sediments to the water between Holmes Road and New Lennox Road during low flow periods is surprising (p.6-72 Model Validation Report) given the otherwise dominant roles of erosion and deposition (Fig 6.2-62 Model Validation Report). I have never seen a system, especially a riverine system, in which diffusion played such an important role. I think that it was necessary to invoke a high diffusion rate to explain the concentration data. I am not sure if there is any precedent for such a high diffusion rate though. This may be perhaps point to another parameterization problem in the model. I recommend the authors investigate this process in more detail and explore other options for calibrating the model.

Response 1B-FG-1:

The Reviewer comments that the PCB "diffusive flux" is too high. Given the context of the use by the Reviewer of the term diffusive flux, the comment is unclear to EPA.

It should be noted that the sediment-water mass transfer flux is often described as a “diffusive flux” because the equation used to compute the mass transfer has the same mathematical form as equations to describe diffusion. The usage of the term “diffusive flux” is unclear in the comment because the mass transfer process for PCBs represents the composite effect of several processes such as diffusion (i.e., molecular diffusion), groundwater advection, and bioirrigation.

As EPA understands this comment, it is possible that the Reviewer used the term “diffusive flux” in the context of molecular diffusion rather than sediment-water mass transfer. In that context, it is worth noting that sediment-water mass transfer coefficients for PCBs are always expected to be orders of magnitude larger than corresponding rates for molecular diffusion because the mass transfer process is a composite that includes processes far larger in scale than molecular diffusion.

If the Reviewer did not misunderstand the usage of the term and “diffusive flux” was meant to indicate the sediment-water mass transfer process, then it should be recognized that the mass transfer coefficient used in the model is well within the bounds values determined from measurements as described in FMD Appendix B.10.

Regardless of the Reviewer’s intended usage of the term “diffusive flux,” the pore water PCB concentrations in Housatonic River sediments are very high, particularly in areas where organic carbon concentrations in the bed are low. These high pore water PCB concentrations are a major determinant of the sediment-water mass transfer flux.

Wilbert Lick

The sediment-water flux of PCBs due to “diffusion” as calculated by the present model is too low. The formulation of this flux and its parameters are also incorrect. The mass transfer coefficient, k, was given a value of 1.5 cm/day, constant throughout the river; this value was chosen on the basis of calibration, not on the basis of any field or laboratory measurements.

Response 1B-WL-1:

EPA disagrees with the Reviewer’s comment that the formulation of the mass transfer coefficient and its parameters is incorrect, and that the value is too low. EPA also disagrees with the characterization that the value was chosen based solely on calibration and not on the basis of field or laboratory measurements. Values for the mass transfer coefficient are well within the range determined strictly from measurements. See the General Response above and Appendix B.10 to the FMD.

Since the PCB flux due to resuspension/deposition should be smaller than that predicted by the present model, the PCB flux due to “diffusion” must increase, strictly on a calibration basis, to compensate for this; it must also increase on the basis of a more fundamental investigation of molecular diffusion and bioturbation (Deane et al. 1999, Lick et al. 2006, Luo et al. 2006). In these articles, it is shown that k due to "diffusion" of HOCs has a lower limit of 1.2 cm/day (no organisms) and increases to 10 cm/day (10^4 oligochaetes/m^2) and even greater for larger numbers of organisms, i.e., the sediment-water flux due to "diffusion" may be significantly greater than that predicted by the present model. A realistic value for k should be determined on the basis of the concentrations and types of benthic organisms. When this is done, k will be significantly larger on the average than it is now and will be relatively low upstream but will increase in the downstream direction.

Response 1B-WL-2:

EPA disagrees with the Reviewer's characterization of the nature of the PCB sediment-water mass transfer process. As demonstrated in FMD Appendix B.10, the mass transfer coefficient for the Housatonic River can be determined from measurements, entirely independent of the model. The mass transfer coefficient value obtained by calibration, K_f = 1.5 cm/day, is well within the range defined by measurements, 1.04 cm/day (geometric mean) to 1.72 cm/day (arithmetic mean). Further, the Reviewer's comment is based on the unfounded assumption that "diffusion" is largely driven by benthic organisms. As discussed in the MCR and MVR and as demonstrated by the sediment profiles images presented in FMD Appendix A.3, physical mixing of river sediment also occurs. See the General Response above.

1.4 GENERAL TOPIC 1C: EQUILIBRIUM PARTITIONING

1.4.1 Comment Summary

Reviewer Name	Page Number/Line Number
Adams	Page 1, line 10 to page 2, line 27
Endicott	Page 4, lines 11 to 20
	Page 4, lines 21 to 30
Gobas	Page 15, lines 1 to 28
Lick	Page 8, line 1 to page 9, line 10
	Page 10, lines 7 to 12
	Page 11, lines 5 to 18
	Page 12, lines 34 to 42
List	Page 1, line 41 to page 2, line 3

1.4.2 General Topic Summary

PCBs are hydrophobic and readily partition between dissolved, bound, and particle-associated (particulate) phases. Partitioning to bound and particulate phases is a function of affinity for surfaces and organic carbon (Karichoff et al., 1979; Schwarzenbach et al., 1993; Chapra, 1997). Mechanistically, partitioning is a function of equilibrium rates at which chemicals sorb (move out of the dissolved phase) and desorb (move back into the dissolved phase). The PCB fate and transport computations in the EFDC model application to the Housatonic River are based on the concept of three-phase equilibrium partitioning (EqP), which assumes that the concentration of freely dissolved PCBs is always in equilibrium with the DOC-bound (complexed) and particulate (sorbed) fractions.

Three Reviewers commented that a time-dependent desorption approach should be applied in the model. One of these three Reviewers commented that it needs to be applied only to the water column and was not necessary for sediment. One Reviewer commented that the EqP approach was appropriate for both water and sediment.

Two Reviewers questioned EPA’s interpretation of the assumption made regarding the relationship between K_{DOC} and K_{POC} and the interpretation of the work performed by Burkhard (2000).

1.4.3 General Topic Response

The use of EqP was reviewed and compared to time-dependent sorption to ensure EqP is appropriate for use in the Housatonic River PCB fate and transport model. Careful consideration was given to the details of how PCB partitioning was implemented in EFDC. In addition to the information summarized in Appendix B.9 of the FMD, PCB partitioning for the modeling study has been described in several other documents, including Attachment B.9 of the Model Calibration Report (WESTON, 2004b), and General Topic 5 of the EPA Response to Questions from the Model Validation Document Overview Meeting (EPA, 2006).

Numerous PCB transport models are described in peer-reviewed literature as well as in project reports, including applications for the Hudson River (Thomann et al., 1991; QEA, 1999; EPA, 2000), the Scheldt River in the Netherlands (Vuksanovic et al., 1996), and the Lower Fox River (Velleux and Endicott, 1994; Steuer et al., 1995; WDNR, 2001). Each of these models (as well as other model applications for sites including Green Bay, Lake Michigan, Saginaw Bay, Milwaukee River, Sheboygan River, etc.) assumes EqP for PCBs. This body of literature demonstrates that EqP is well established as the state of the practice for large-scale PCB transport model applications. The EFDC application for the Housatonic River PSA is fully consistent with this established state of the practice. Further, no example of a large-scale PCB transport model that uses a kinetic (non-equilibrium) approach was found in the peer-reviewed literature. While this does not mean such kinetic applications do not exist, it does demonstrate that all practical models in the literature utilize the EqP approach.

The EFDC application to the Housatonic River uses EqP for PCB transport and fate calculations. This approach is consistent with the state of the practice for large-scale PCB model applications. For most circumstances, the conditions for use of EqP are satisfied. During high-flow events that resuspend sediment from the riverbed, it is possible non-equilibrium conditions may exist. However, even for such high-flow events, the maximum expected difference between water column phase distributions computed using EqP or kinetic approaches will be approximately 5 to 10%. Therefore, EqP provides a reasonable basis for modeling PCB fate and transport within the Housatonic River.

1.4.4 Response to Specific Comments on General Topic 1C

Eric Adams

The model assumes that three chemical phases of PCBs are in equilibrium. As EPA points out, this is a common modeling assumption, but I believe it is only reasonable within a stationary sediment bed, and not within the water column during sediment resuspension.

Following Wu and Gschwend (1986), and focusing only on sorption and desorption among two phases (sorbed PCB with concentration C_s and dissolved PCB with concentration C_d)

$$\begin{aligned}\frac{dC_s}{dt} &= k_1 K_p C_d - k_1 C_s \\ \frac{dC_d}{dt} &= k_2 C_s / K_p - k_2 C_d\end{aligned} \qquad (1)$$

where k_1 and k_2 are rate constants constrained by

$$k_1 = k_2 / \rho K_p \qquad (2)$$

where ρ is the solid-water phase ratio $\rho = (1-\phi)\rho_s / \phi$, ρ_s is the solid density, ϕ is porosity and K_p is the partition coefficient. For a stationary sediment bed, ρ is about 1, so for hydrophobic contaminants (large K_p), most of the contaminant is sorbed; hence the processes of sorption and desorption cause C_d to vary in the range $0 \le C_d \le C_s / K_p$, whereas C_s remains nearly constant. The characteristic time for sorption/desorption for highly particle-reactive species in a stationary sediment bed ($\rho K_p >> 1$) is thus k_2^{-1}.

Wu and Gschwend (1986) describe the sorption/desorption process as one of molecular diffusion through particles (or particle aggregates). For large K_p and an assumed intra-aggregate porosity, their effective (retarded) diffusion coefficient is given by

$$D_{eff} \cong \frac{0.2 D_m}{K_p \rho_s} \qquad (3)$$

where D_m is the molecular diffusivity of PCBs. Wu and Gschwend (1988) fit a first-order sorption kinetics model to the radial diffusion model over the (initial) time during which half of the sorption/desorption takes place. Comparing the two models, and for $\rho K_p >> 1$

$$k_2^{-1} \sim \frac{(R / \rho K_p)^2}{D_{eff}} \qquad (4)$$

where R is a characteristic aggregate radius. It can be seen that the time scale (k_2^{-1}) is proportional to K_p^{-1}. In other words, for highly particle reactive species, little of the initially sorbed mass is exchanged during desorption and that which is, is lost from the outside of the particle (over an effective length of order $R/\rho K_p$). Even for a large R (0.01 cm) and $K_p = 10^5$ cm^3/g, the time scale (k_2^{-1}) is less than one second, suggesting that desorption is very fast, and that the assumption of ***equilibrium partitioning is probably acceptable within the sediment bed***. See also following discussion under diffusive exchange.

Response 1C-EA-1:

EPA appreciates the detailed analyses provided by the Reviewer. As noted by the Reviewer, average contact times in the sediment bed are typically so long that the use of EqP is appropriate. See General Response above.

During resuspension, where the sediment concentration in the water column (TSS) is small, equilibrium partitioning would require that most contamination become desorbed. Hence diffusion must take place through the entire particle, giving a time scale (k_1^{-1})

$$k_1^{-1} \sim \frac{R^2}{D_{eff}} \qquad (5)$$

Even for a small R (0.001 cm), the time scale is over a day, which exceeds the duration during which most suspended particles remain in the water column. Hence the assumption of ***equilibrium partitioning is not really valid for resuspended particles***. Because much of the contamination on resuspended particles will not have time to de-sorb, assuming equilibrium partitioning overestimates the PCB flux from bottom sediments due to resuspension. Of course, during calibration, this could have been compensated for, in part, by assuming less sediment resuspension.

Response 1C-EA-2:

EPA disagrees with the Reviewer’s comment regarding the use of EqP for resuspended particles on the basis of numerical tests of the Housatonic River model conducted with EFDC. Those tests are presented in FMD Appendix B.9 and document that the expected difference in PCB phase distributions (i.e., the fraction of the total PCB that is dissolved or particulate) computed for a short resuspension event using EqP or kinetic approaches will be approximately 5 to 10%. This difference is minor given the large uncertainty associated with PCB adsorption and desorption rate measurements. Based on this analysis, EPA believes that EqP is an appropriate representation of the PCB phase distribution in the water column as well as the sediment bed. See General Response above.

Douglas Endicott

EFDC uses the same equilibrium partitioning (EP) model as every other fate and transport model in use today. EP generally works quite well, and simplifies the model computation by requiring only a single state variable for each chemical. It has been suggested that desorption kinetics should be incorporated into the Housatonic River model, because PCBs desorbing from resuspended sediment might not reach equilibrium within the timeframe they remain in suspension. This disequilibrium may help explain the pattern of PCBs moving in "ribbons" of sediment downstream, and the high surficial PCB concentrations in Woods Pond sediment. I think this is mostly speculation and, although I am curious to see if it is true, it is well and beyond the objectives and goals of this modeling study.

Response 1C-DE-1:

EPA agrees with the Reviewer's comment that EqP is the state of the practice for fate and transport models in use today. EPA investigated the use of time-dependent desorption in the model and corresponding model sensitivity. Further, the difference between EqP and kinetic partitioning approaches is expected to be approximately 5 to 10%, as documented in FMD Appendix B.9.

I have already commented on the misattribution of the assumption made regarding PCB partitioning to non-filterable organic carbon in the water column ($K_{doc} = K_{poc}/100$), and EPA has responded (less than satisfactorily) to this issue. Regarding the parameterization of the 3-phase equilibrium partitioning model for PCBs, a useful citation (i.e., the Burkhard references) is: USEPA, 2003. *Methodology for Deriving Ambient Water Quality Criteria for the Protection of Human Health (2000), Technical Support Document Volume 2: Development of National Bioaccumulation Factors*. United States Environmental Protection Agency, Office of Water, Office of Science and Technology. EPA-822-R-03-03
(www.epa.gov/waterscience/criteria/humanhealth/method/tsdvol2.pdf)

Response 1C-DE-2:

The modeling team has again considered this comment at length. The modeling team still remains unsure as to the exact technical details and scope of the Reviewer's comment. As noted by the Reviewer, this topic was first raised at an earlier stage of this Peer Review. At that time, the modeling team believed it had fully responded to the Reviewer's comment. Specifically, the Reviewer repeated this comment and noted that he found the modeling team's response "less than satisfactory". However, it must be noted the EPA had made every effort to fully address the Reviewer's comment as it was understood.

Nonetheless, in an effort to again resolve this comment, further descriptions of EPA's technical basis for selecting PCB partitioning parameters for the Housatonic River model framework are provided below.

As part of model development, the modeling team assumed that $K_{POC} \approx K_{OW}$ as described by DiToro (1985). From this assumption of the approximate equality of K_{POC} and K_{OW}, it is concluded that $K_{DOC} = \alpha_{DOC} K_{POC} \approx \alpha_{DOC} K_{OW}$, as noted by Burkhard (2000). With respect to the DOC binding efficiency (α_{DOC}), it is worth noting that for sediment $\alpha_{DOC} = 0.1$, while for the water column $\alpha_{DOC} = 0.01$. Both the values for α_{DOC} for sediment and α_{DOC} for the water column were determined from site-specific data for the Housatonic River (see FMD Appendix B.9).

DOC binding effectiveness in the water column will typically be less than in the sediment bed, as described by Eadie et al. (1990, 1992). One explanation for the decreased binding efficiency in surface waters is photobleaching of DOC by ultraviolet (UV-B) radiation (Kashian et al., 2004). On the basis of these references and site-specific partitioning data, EPA believes that the relationship between K_{POC} and K_{DOC} is appropriately defined for the Housatonic River.

Frank Gobas

The model assumption that the sorptive capacity of DOC is two orders of magnitude less than that of POC (l.27-28, p.6-67 Model Validation Report) seems out of touch with the literature. The lowest value I have seen to characterize the sorptive capacity of DOC compared to octanol was 0.08, i.e. DOC has 8% of the sorptive capacity of octanol (Burkhard, L.P. Estimating dissolved organic carbon partition coefficients for nonionic organic chemicals. *Environ. Sci. Technol.* 2000, *34* (22), 4663-4668. – Note that in response to p. 2-72 of EPA response, Burkhard states that K_{DOC} = 0.08.Kow, not K_{DOC} = 0.08.K_{POC}). In comparison, the sorptive capacity of POC is approximately 35% of that of octanol (Seth, R.; Mackay, D.; Muncke, J. Estimating the organic carbon partition coefficient and its variability for hydrophobic chemicals. *Environ. Sci. Technol.* 1999, 33 (14), 2390-2394.). Following these papers, the difference in sorptive capacities between DOC and POC is approximately a factor of 35/8 or 4.38. Burkhard et al. did report 95% uncertainty limits of a factor of 20 for the 0.08 value, hence the 0.01 value used in the model is plausible, but it is very low.

Much higher values of the sorptive capacities of DOC in relation to octanol have also been measured. For example, Macintosh et al. report 1.16 (± 0.49) for spiked and 61 (± 47) for native PCBs and phthalates. This compared to 35 (± 19) for POC (Sorption of Phthalate Esters and PCBs in a Marine Ecosystem, Mackintosh et al. Environ. Sci. Technol.; 2006; 40(11); 3481-3488.). The latter results indicate no significant differences between sorptive capacities of DOC and POC and also evidence of DOC-water disequilibria.

The assumed two orders of magnitude difference in sorptive capacities of POC and DOC in the model is, albeit plausible, a very low value. Given the variation in literature data, I recommend that empirical data are used to calibrate this model input requirement. The recent EPA response document suggests that the latter has indeed been done.

Response 1C-FG-1:

As noted by the Reviewer, the sorptive capacity of POC relative to that of octanol is variable. However, DiToro et al. (1985) found that $K_{POC} \approx K_{OW}$. Based on the assumption of approximate equality of K_{POC} and K_{OW}, it is concluded that $K_{DOC} = \alpha_{DOC} K_{POC} \approx \alpha_{DOC} K_{OW}$, as noted by Burkhard (2000). With respect to the DOC binding efficiency (α_{DOC}), it is worth noting that for sediment $\alpha_{DOC} = 0.1$, while for the water column $\alpha_{DOC} = 0.01$. Both the values for α_{DOC} for sediment and α_{DOC} for the water column were determined from site-specific data for the Housatonic River (see FMD Appendix B.9).

> However, it should be noted that DOC binding effectiveness in the water column will typically be less than found in the sediment bed, as described by Eadie et al. (1990, 1992). One explanation for the decreased binding efficiency in surface waters is photobleaching of DOC by ultraviolet (UV-B) radiation (Kashian et al., 2004). On the basis on these references and site-specific partitioning data, EPA believes that the relationship between K_{POC} and K_{DOC} is appropriately defined for the Housatonic River.

Wilbert Lick

After sediment particles are resuspended, they will be transported downstream by the current and eventually settle out of the water column. During this time, the contaminant sorbed to the particle will desorb at some finite rate. The time for a particle to settle out of the water column depends on the settling speed and water depth, while the distance traveled by the particle before depositing depends on the settling time and current speed. For a reasonable range of settling speeds, w, for fine-grained particles/flocs (2×10^{-3} to 1×10^{-1} cm/s) and water depths, h, typical of the Housatonic (1 to 3 m), the settling times ($t = h/w$) are in the range (see table) from 10^3 s (15 min) to 1.5×10^5 s (1.5 days). For medium and coarse grained particles, the settling times are less.

Settling Times (Seconds) for Fine-Grained Particles in the Housatonic

h(cm)	w(cm/s)	
	2×10^{-3}	1×10^{-1}
1×10^2	0.5×10^5 (0.5 days)	1×10^3 (15 min)
3×10^2	1.5×10^5 (1.5 days)	3×10^3 (45 min)

Some experimental results for the adsorption and desorption of HOCs are shown in the appended figures. These sorption times depend on the sediment concentration, particle and floc sizes, conditions of the experiment, and the value of the partition coefficient. The first three figures are for hexachlorobenzene ($K_p = 10^4$ L/kg), while the fourth figure is for the adsorption of HOCs with K_p's from 10^3 to 6.6×10^4 L/kg. The last figure is for the desorption of a PCB with one chlorine (MCB, $K_p = 10^3$ L/kg), hexachlorobenzene (HCB, $K_p = 10^4$ L/kg), and a PCB with six chlorines (HPCB, $K_p = 6.6 \times 10^4$ L/kg, a K_p which is smaller than, but comparable to, the average K_p of about 10^5 L/kg for PCBs in the Housatonic). In 10 days, only about 25% of the HPCB has desorbed; in 50 days, only about 55% has desorbed. As the partition coefficient increases, the amount of desorption in these time intervals will be even less. For PCBs with $K_p = 10^5$ L/kg, the desorption times would be approximately 1.5 times greater than those for HPCB shown here.

Since desorption times are much greater than settling times, it follows that contaminants on resuspended particles will not desorb completely, or even close to completely, in the water column before the particles settle out of the water column. The chemical sorbed to the suspended particles will therefore not reach chemical equilibrium with the chemical dissolved in the water column. It follows that the assumption of equilibrium partitioning is not valid, nor even a good approximation, for the sediments in suspension or in the surficial layers of the bottom sediments.

Incidentally, finite rates of adsorption/desorption (1) will assist in explaining the high observed values of PCBs in the surficial sediments of Woods Pond, (2) the high variability in observed PCB concentrations in the sediments, and (3) probably will force a higher n in erosion formulas (consistent with all other data) since non-equilibrium sorption will not be consistent with EPA's imposed constraint due to equilibrium partitioning.

The conclusion is that finite rates of PCB adsorption and desorption have a major effect on the sediment-water flux due to resuspension/deposition and must therefore be considered in the modeling.

Response 1C-WL-1:

See General Response above and Appendix B.9 to the FMD. PCB phase distributions computed using EqP and kinetic approaches as suggested by the Reviewer are expected to differ by just 5 to 10% in the Housatonic River. This difference is minor given the large uncertainty associated with PCB adsorption and desorption rate measurements. It should be noted that the issue is whether EqP is an adequate approximation for the model and not whether PCB adsorption or desorption processes have reached full thermodynamic equilibrium. Based on this analysis, EPA believes that EqP is an appropriate representation of the PCB phase distribution in the water column.

EPA strongly disagrees with the Reviewer's conclusion that the difference in PCB phase distributions as computed using EqP and kinetic approaches is "major." For the Housatonic River, this difference is expected to be approximately 5 to 10%.

Since equilibrium partitioning is not valid, a new parameter (the sorption rate) is introduced into the problem. Among other things, this invalidates EPA's statement (when speaking of non-unique solutions) that "there is a constraint imposed by PCB transport that results from resuspension and deposition processes." This constraint, if it exists, is incorrect because non-equilibrium sorption would impose an entirely different constraint than that imposed by equilibrium partitioning.

Response 1C-WL-2:

As noted in the General Response and Response 1C-WL-1, EqP is an appropriate representation of PCB partitioning in the Housatonic River.

In my comments above, I emphasized that PCB sorption times are slow relative to particle settling times and that, because of this, equilibrium partitioning (as assumed in the model) was not a good assumption. Consider the effects of this on PCB transport to Woods Pond. The upstream region of the Housatonic is erosional (on the average), and PCB concentrations near the sediment surface are relatively high (because they were deposited at an earlier time before remediation). As these sediments are eroded, PCBs tend to desorb from the suspended sediments but do not reach anywhere near chemical equilibrium before the sediments are deposited, i.e., the PCB concentrations of depositing sediments are much higher than if they had equilibrated in the overlying water. This erosion/deposition may occur several times before the sediments and their sorbed PCBs reach Woods Pond, or it may only occur once, depending on the flow rate and turbulence. Either way, the sediments deposited in Woods Pond will have higher PCB concentrations than if equilibrium partitioning was assumed, as is observed in field measurements but is not predicted by the present model.

Response 1C-WL-3:

The reviewer's assumption that EqP is not appropriate for the Housatonic River is incorrect. See FMD Appendix B.9. EqP adequately simulates the expected phase distribution for PCBs.

Equilibrium partitioning is not a valid assumption. Desorption rates are relatively slow such that, when bottom sediments are resuspended, the sorbed PCBs do not desorb sufficiently rapidly for equilibrium partitioning to be approached before particle deposition occurs. The result is that PCBs sorbed to the suspended solids are not in equilibrium with the PCBs dissolved in the water. The sediment-water flux of PCBs due to resuspension/deposition is therefore much lower than that predicted by equilibrium partitioning, as is assumed in the present model. A finite sorption rate between the PCBs sorbed to the solids and the dissolved PCBs needs to be added to the model and will replace the equilibrium assumption.

Response 1C-WL-4:

EPA disagrees with the Reviewer's assertion that EqP is not valid. See General Response above and FMD Appendix B.9. The difference in phase distributions computed using EqP and kinetic approaches is approximately 5 to 10%.

E. John List

Because the PCBs are an integral part of the stream sediments the difficulties of TSS mass balances carry through to the PCBs. To compound the difficulty, the PCB mass flux has two components: one associated with the sediments and the other with the dissolved phase in the water. A further complication is the kinetics of the partitioning process of the PCB between the water and the sediment, which was a matter of some discussion at the Peer Review meeting and may need addressing.

Response 1C-JL-1:

EPA has evaluated the kinetics of the partitioning process (see General Response above).

1.5 GENERAL TOPIC 1D: BANK EROSION

1.5.1 Comment Summary

Reviewer Name	Page Number/Line Number
Bohlen	Page 8, lines 16 to 23
Garcia	Page 1, line 39 to page 2, line 18
	Page 2, line 37 to page 3, line 6
List	Page 1, lines 39 to 41

1.5.2 General Topic Summary

The model framework includes bank erosion because of its importance to longer-term simulations as identified in the MFD and Reviewers' comments from the previous two Peer Reviews.

Two Reviewers commented on the adequacy of the formulation used to represent bank erosion in EFDC. One Reviewer noted that bank erosion can influence channel morphology.

1.5.3 General Topic Response

Bank erosion has two main components (Knighton, 1998): toe erosion by hydraulic action (fluvial erosion) and mass failure. The total bank erosion rate is the sum of these two components and can be estimated from field data using a range of approaches. Separating total erosion rates into toe erosion and mass failure components is difficult without direct time series observations and is still an area of active research. Site-specific data for Reaches 5 and 6 define short-term and long-term average rates of bank erosion but do not define the time history of erosion.

For the Housatonic River EFDC model application, an empirical approach was used to estimate bank erosion rates. Long- and short-term bank erosion solids loads were determined from field investigations, including the differences between geo-referenced aerial photographs and survey data (Woodlot, 2002, which is included as Appendix A.4 to the FMD).

Although the exact time history of bank erosion cannot be determined directly from these surveys, the overall bank erosion loads used in the model are constrained by the bank erosion solids mass entering the river as determined from site-specific data for long-term (1972–2001) and short-term (2001-2002) periods.

To approximate the time history of bank erosion, two types of solids loads were input to EFDC: (1) continuous loads that represent erosion of the bank toe below the water line; and (2) periodic loads that represent mass failure for an event. Both loads were computed as empirical functions of flow, with the total bank erosion defined as the sum of toe erosion and mass failure components.

The empirical functions developed to estimate bank erosion loads for EFDC model application to the Housatonic River are consistent with long-term and short-term bank erosion rates determined from field data and site-specific analyses. As parameterized, the empirical equations yield bank erosion loads that match the long-term and short-term erosion rates with a very small residual error. While there is uncertainty associated with the time history of these loads, use of a mechanistic model to generate an alternative time history would not eliminate or reduce this uncertainty. Further, it should be recognized that bank erosion represents an appreciable but nonetheless small component (13%) of the annual average external solids budget for the river. Given the limitations and variability associated with bank erosion measurements for the site, and because it is a small component of the overall solids budget, further refinement of the bank erosion time history is neither practical nor necessary to achieve the goal of the modeling study. See Appendix B.7 to the FMD.

1.5.4 Response to Specific Comments on General Topic 1D

W. Frank Bohlen

Beyond this matter of flow, velocity and shear stress, I was pleased to see that the current model includes direct calculation of side bank erosion. I'll leave the adequacy of this formulation to those more qualified than I but must state my concern over the apparently simple partitioning of the mass of sediment supplied by this process between surficial erosion (during the rising hydrograph) and mass failure (during falling stage). It's hard for me to see why the masses should in anyway be equivalent. This subject was also noted in GE's review of the Model

Validation Report (MVR). Justification for this approach should be carefully presented using the field data and/or supplementary data from previous publications.

Response 1D-FB-1:

See the General Response above. Estimates of bank erosion for the Housatonic River are fully consistent with site-specific bank erosion data.

Marcelo H. Garcia

It is not enough to state that a given process is accounted for when the representation of the process is deficient and its implementation in the model is done with algorithms and assumptions that are not based on the physics of the process and cannot be supported by field observations. For instance, streambank erosion has been recognized as a very important process since large amounts of PCB-laden sediments can enter the river during medium to high flows. However, fluvial erosion and mass failure of the stream banks are currently modeled simply as functions of flow discharge, a rather crude approximation, which severely limits the capabilities of the model to assess the effect of this process.

Response 1D-MG-1:

EPA disagrees with the Reviewer's comment. Bank erosion estimates in the model are based on empirical functions that accurately reproduce the mass of solids eroded from the river banks over time. Bank erosion is a complex topic in general and the exact mechanism driving bank erosion in the Housatonic River system cannot be determined from the present data. However, the empirical functions developed for the model are based on information describing river bank erosion conditions for the past 3 to 5 decades. Although the reviewer may regard the empirical approach as "crude," it should be recognized that it is accurate. See General Response above and FMD Appendix B.7.

How can one determine what regions of the banks are more prone to erosion and, therefore, need protection if the model does not compute the distribution of flow velocity and bed shear stress along the banks? Could native vegetation or other bioengineering techniques be used to protect the banks of the Housatonic River against erosion or more hard-core solutions such as rip-rap are needed? Is it possible to use a combination of both? As it stands now, the model can not be used to compare the merits of such remedial alternatives, which is one of the goals of the modeling study as stated above.

Response 1D-MG-2:

As EPA has stated to the Reviewers, it is not the goal of the modeling study to develop models that will be used to evaluate specific engineering applications on a location-specific basis in the CMS. The goal of the model (in the situation described by the Reviewer) is merely to assess if the contribution of PCBs from the banks in given reaches warrants attention when evaluating remedial alternatives, and the relative impact of the bank contributions on alternatives. The level of detail of modeling needed to evaluate this contribution for a remedial alternative would be to simply "control" bank erosion in the model by setting the bank source contribution to zero as was done by EPA in one of the demonstration runs.

The Reviewer, however, questions the capabilities of the model for just such an application. The model will not be used to evaluate the specific areas of bank that will be subject to a specific engineering control such as bioengineering or hard engineering. There are other tools specifically developed to assist in evaluating these engineering techniques. Such decisions would be made in the design phase of a remedy, if necessary.

At the same time, the use of a course grid hinders the ability of the model to resolve the flow velocity distribution inside the main channel as well as the associated boundary shear stresses throughout its wetted perimeter. Given that erosion and sediment transport vary non-linearly with increments in flow velocity and shear stress, it is rather imperative to find a way to increase the spatial resolution of the model so that at least three "streamtubes" are used to model flow and sediment transport in the main channel of the Housatonic River. This will in turn facilitate and make more meaningful the computation of streambank erosion as well bed sediment erosion and resuspension.

Response 1D-MG-3:

See General Response. The empirical functions used to estimate bank erosion are based on flow, not velocity, and very accurately represent measured long-term and short-term erosion rates in the Housatonic River. See FMD Appendix B.7.

E. John List

If, in addition, the stream bank erodes then long term changes in the direction and shape of the channel can result.

Response 1D-JL-1:

The Reviewer correctly notes the linkage between bank erosion and channel morphology. See General Response above.

2. GENERAL TOPIC 2: SPATIAL TRENDS AND MODEL-DATA COMPARISONS

2.1 GENERAL TOPIC 2A: PREDICTED SEDIMENT/WATER COLUMN PCB CONCENTRATIONS

2.1.1 Comment Summary

Reviewer Name	Page Number/Line Number
Adams	Page 7, line 25 to page 8, line 10
Bohlen	Page 12, lines 14 to 25
Endicott	Page 7, lines 8 to 20
	Page 7, line 25 to page 8, line 5
	Page 10, lines 9 to 16
Garcia	Page 3, line 35 to page 4, line 16
	Page 5, lines 17 to 25
Gobas	Page 2, line 43 to page 3, line 30
	Page 3, lines 39 to 43
	Page 6, lines 4 to 24
	Page 7, lines 19 to 31
	Page 7, line 44 to page 8, line 2
	Page 8, lines 10 to 32
	Page 13, lines 9 to 25
Lick	Page 1, lines 30 to 33
	Page 10, line 38 to page 11, line 3
	Page 13, lines 21 to 26
List	Page 4, lines 4 to 8
	Page 4, lines 18 to 21
	Page 5, lines 5 to 25

2.1.2 General Topic Summary

The identification of trends in environmental media is important in assessing model performance because one of the most common ways of evaluating model performance is to compare the simulated concentrations in sediment and fish tissue with those measured. However, accurate quantification of spatial and temporal PCB concentration trends is complex. The majority of the sampling programs for the Housatonic River were not designed to estimate PCB concentration trends over time. Differences in sediment PCB concentrations resulting from spatial heterogeneity, temporal variability, sample size, and analytical bias confound analysis and make clear identification of trends challenging. It is important to carefully evaluate the occurrence of significant trends in contaminant concentrations using rigorous statistical methods that take into account these potentially confounding factors. Often an apparent trend (one presumed by viewing the data) is in fact not significant (does not exist in a statistically meaningful way).

Changes in surface water, sediment, and fish tissue PCB concentrations over time were initially described in the RFI (BBL and QEA, 2003), and EPA documents such as Responses to Questions from the Reviewers. Statistically significant changes in concentrations are now more fully quantified in FMD Appendix A.1.

All Reviewers commented on the ability of the model to simulate concentrations of PCBs in surface water and sediment, including indications in the data of real or apparent spatial patterns and/or temporal trends in PCB concentrations in the PSA. Four of the Reviewers combined their comments on these model-data comparisons with additional comments about small-scale spatial variability in PCB concentrations, which is addressed in General Topic 2B.

All Reviewers commented that they believed there is, or should be, a trend of decreasing PCB concentrations in Woods Pond sediment, yet the model did not reproduce the magnitude of this trend and was thereby biased high with respect to the amount of time necessary for natural remediation to reduce sediment PCB concentrations in Woods Pond. One Reviewer commented further that even if the model was biased in this regard, it was apparent that any such natural remediation would require “decades”; therefore, the question of bias may be moot.

Most Reviewers recommended additional study to better understand the behavior of the model regarding spatial and temporal trends. One Reviewer recommended a more refined statistical study of the data to better characterize the perceived temporal trend in Woods Pond sediment PCB concentrations.

2.1.3 General Topic Response

In response to Reviewer comments, EPA performed further data analyses of trends in surface water, surface sediment, and fish tissue PCB concentrations for Reaches 5 and 6 beyond those presented previously in the RFI and other documents. These analyses are summarized in Appendix A.1 of the FMD.

Analyses indicate that the trend (or decline) in surface sediment PCB concentrations in Reaches 5 and 6 is not statistically different from zero (i.e., slow to no change over time). Model results are consistent with this observation. Specifically, the simulated rate of surface sediment PCB concentration decline in EFDC is slow and within the uncertainty envelope of the data. There is no statistically significant trend in PCB concentrations in Woods Pond surficial sediment. In contrast, the Reviewers assumed that there was a more rapid decline in sediment PCB concentrations over time and then incorrectly concluded that the rate of decline in PCB concentrations simulated by the model (which is in agreement with the data) was too slow. The Reviewers may have based their assumption on what could appear visually to be a decline in Woods Pond sediment. However, this decline is not statistically significant because of the limitation of sample size available for certain time periods, and an analytical bias (low) in much of the sediment data collected in the later portion of the time period. This analytical bias was discussed in Attachment 4 to the MCR, and in Appendices A.1 and A.2 to the FMD.

Statistical analyses of trends in surface water concentrations for the validation period were conducted and are documented in Appendix A.1 of the FMD. These analyses indicate a statistically significant decline in PCB concentrations ranging from 5.8 to 13.8% per year in Reaches 5 and 6 for the period from 1979 to 2004. It is not meaningful to estimate a trend in PCB concentrations in surface water from EFDC results because the EFDC-simulated surface water concentrations are sensitive to the boundary condition function, which was essentially constant over the 26-year period. Within that time, EFDC boundary conditions would be influenced solely by the hydrograph, and would not necessarily reflect the historical concentrations in loads that predate the data set used to develop the upstream boundary flux function.

The analysis of spatial and temporal trends in PCB concentrations in fish tissue shows a slight decline over time in whole-body fish tissue PCB concentrations in Reaches 5 and 6. The decline in fish tissue concentrations was statistically significant for some of the species and/or subreaches and not for others. The magnitude of the decline was generally small in most cases, with a half-life on the order of decades. This is consistent with the results of the analysis of trends in sediment PCB concentrations. FCM accurately simulates the concentrations of PCBs in fish tissue, with a simulated rate of decline that is slow and within the uncertainty envelope of the data.

2.1.4 Responses to Specific Comments on General Topic 2A

Eric Adams

The spatial plots at low Q (Figure 4.2-69) and moderate Q (Figure 4.2-70) show that upstream concentrations of both TSS and PCB are too low (a consequence of the BC) and that there is more local variability (indicated by shading) than variability between reaches. There is no way to verify the local variability. It is hard to say whether the overall spatial trend is correct or not, because of variability in the data. The temporal trends look OK.

Response 2A-EA-1:

The Reviewer is correct in noting that the variability in the data complicates evaluation of spatial and temporal trends. The figures identified by the Reviewer in his comments regarding spatial patterns present model-data comparisons for low-flow conditions (flows at Coltsville between 16 and 25 cfs) and conditions slightly above average flow conditions (flows at Coltsville between 109 and 275 cfs). These figures highlight the effect of the spatial transient in PCB concentrations due to PCB mass transfer from the sediment under low- to moderate-flow conditions. An additional and perhaps more important demonstration of model performance is the ability of the model to reproduce the roughly one order of magnitude decline in water-column PCB concentrations and also the decline is TSS concentrations between New Lenox Road and Woods Pond Outlet under high-flow conditions (Figures 3.5-6 and 3.5-15 in the FMD, previously presented as Figures 4.2-75 and 4.2-76 in the MVR).

The temporal changes in sediment PCB (top 6 inches) show that the model does not reproduce the significant decline in Woods Pond over time (Figure 6.2-50). This is one of GE’s main points and suggests that the model might also underpredict the response to remedial actions, including natural attenuation. ***I see this as a major concern***.

The model misses the decline in PCBs across Woods Pond that is both observed and intuitively suspected. This would mean the model overpredicts loading downstream (downstream model).

Response 2A-EA-2:

> The statistical analysis of Woods Pond sediment PCB data, described in Appendix A.1 of the FMD, indicates that there is no statistically significant change in surface sediment concentrations in Woods Pond over time as the Reviewer incorrectly assumes. The simulated rate of sediment PCB concentration decline in EFDC is slow and is well within the uncertainty envelope of the data. This more rigorous statistical analysis supports the characterization of model-data comparison of PCB concentrations discussed in Section 6.2.5.1 of the MVR and in previous documents, including the RFI.

Like TSS, the equivalence plots for PCBs (Figure 4.2-85) shows less variability than the data. However, there is little indication of (overall) bias.

And also like TSS, it would also be nice to have had more data on PCBs. Again, there is better model-data agreement at relatively high (presumably more biologically relevant) concentrations. As with other variables, the model shows less variability than the data (i.e., it underpredicts the highs and overpredicts the lows).

Response 2A-EA-3:

> EPA has consistently acknowledged in the MCR and MVR that the model results have less variability than the data, and discussed reasons why this is expected. EPA also agrees with the Reviewer that there is better model-data agreement at relatively high PCB concentrations, and that these conditions are more biologically relevant. A bounding analysis (MVR Section 6.2.3.4 and the EPA Response to Questions from the Model Validation Document Overview Meeting) demonstrated that the overpredictions of low PCB concentrations had less than a 5% effect on fish tissue concentrations, averaged across all species.

Except for the transducers which record flow (through a stage discharge relationship), better data seem to be available for the calibration (especially the original calibration) period, than for the validation period. However, I am less concerned with traditional model validation (the model's ability to predict absolute concentrations under different hydrologic conditions) than I am with the model's ability to distinguish among different remedial alternatives that would presumably be evaluated under the same hydrologic conditions. See comments under Question 6.

Response 2A-EA-4:

> The example simulations, which were recommended by this Reviewer, were viewed by the modeling team as an effective way to demonstrate the behavior of the model under very different conditions than those simulated during the calibration and validation periods. Although not intended to be viewed as an evaluation of any remedial alternative, the results of the example simulations provided the modeling team and several Reviewers with additional confidence in the ability of the model to describe (1) the response of the system to changes associated with potential remedial alternatives, and (2) the relative difference simulated in the two alternatives. The "tracer," or source-specific simulations,

provide additional insight into the relative contributions of different processes in different locations to the overall response of the system.

W. Frank Bohlen

In terms of PCBs, the model appears to overpredict concentrations in Woods Pond (see Fig.6.2-50). This is likely the result of inaccurate specification of suspended sediment flux associated with the above TSS predictions. Alternatively it may be associated with the specification of boundary conditions and/or the model treatment of PCB sequestering in Woods Pond. The model decline within the Pond appears to be approximately linear with time while the data suggest an exponential distribution. The latter might be the result of progressive source control (both natural due to burial and anthpogenic) acting in combination with sedimentation of cleaner materials. The model might easily obscure this process since it assumes a well mixed sediment column to 6in. With sedimentation rates in Woods Pond averaging approximately 2-3 mm/yr the above depositional processes might be expected to influence the upper 3in over 25 years. This might very well result in a reduction in actual PCB concentrations that would not be well simulated by the model.

Response 2A-FB-1:

The Reviewer's interpretation of the comparison of trends in sediment PCBs in Woods Pond is erroneous, and contradicted by the results of the rigorous statistical analysis (FMD, Appendix A.1), which supports the interpretation provided by the modeling team in the MVR and Document Overview Meeting presentations. Because the Reviewer's premise is incorrect, no response can be provided for the hypotheses offered by the Reviewer to explain the supposed inconsistency.

It is noted, however, that the Reviewer's description of a 6-inch, well-mixed layer is inconsistent with the model's parameterization of the bed (see FMD Appendix B.5, Tables B.5-2 and B.5-3), which includes a 1.6-inch (4-cm) mixed layer in Reach 5A and a 2.8-inch (7-cm) mixed layer in the remainder of Reach 5 and all of Reaches 6 through 8.

Douglas Endicott

Model predictions of total PCB concentrations in the Housatonic River water column were compared to data using time series plots and longitudinal profiles. The time series plots (Figures 4.2-65 thru 4.2-68 and 6.2-41 thru 6.2-44) indicate that, for PCB concentrations, there are significant gaps in data collection over the Phase 2 calibration and validation periods. As was the case for TSS, the model tends to overpredict low PCB concentrations, and the overprediction of these two state variables appear to coincide. This is most apparent at the Holmes Road station. Again, this is primarily due to significant bias in the specification of the East Branch boundary condition on days when PCB concentrations were not measured. Inspection of the timeseries of East River PCB boundary conditions (Figure 4.2-64) reveals that the only time East River PCB boundary concentrations are lower than 60 ug/L is when data have been measured below this value, yet probably more than half of the data fall below this concentration. There are sufficient PCB data to evaluate the storm event predictions on only one occasion (May 19-21, 1999 at New Lenox Road; Figure 4-2.76); the model prediction timeseries agrees well with this data. The spatial profiles for PCB water column predictions appear to be quite variable.

Response 2A-DE-1:

Solids and PCB concentrations measured at the upstream boundary are highly variable. Across narrow ranges of flow, these measured values can vary by more than two orders of magnitude. For all practical purposes, a many-to-one relationship exists between measured concentration and flow. The high degree of variability in measured East Branch boundary TSS and PCB concentrations within narrow flow ranges presented a significant challenge for the development of estimates of boundary concentrations for times when data were not available. The uncertainty associated with boundary loadings estimates of PCBs has been acknowledged. Appendix B.2 of the FMD provides a discussion of the approaches evaluated to develop boundary conditions, including those proposed by Reviewers. Although none of the techniques considered were able to reproduce the variability of the measured concentrations, the technique that was developed was judged to be acceptable because it yielded estimates that fit the data somewhat better than prior estimation methods did, and a bounding analysis (MVR Section 6.2.3.4, and the EPA Response to Questions from the Model Validation Document Overview Meeting) demonstrated that the overpredictions of the low PCB concentrations had less than a 5% effect on fish tissue concentrations, averaged across all species.

In the *EPA Response to Questions from Model Validation Document Overview Meeting*, a statistical test (overlapping 95% confidence limits) is offered to show that the apparent decrease in PCB concentrations across Woods Pond may reflect sampling variability. It may be useful to use a statistical test to evaluate the gradient in PCB concentration data between two river locations; however, wouldn't it be better (more powerful) to test whether the *difference* in PCB concentrations between locations is significantly different than zero?

Timeseries plots comparing the predictions of PCB concentrations in the top 6" of sediment in mile reaches to data are shown in Figure 4.2-77 for the Phase 2 calibration period, and Figures 6.2-49 for the validation period. The predictions appear to exceed the majority of data at the end of both the calibration and validation periods, in some or most mile reaches. Although data prior to 1999 are not abundant, the model also appears to underpredict the rate of PCB sediment decline in many mile reaches, including Woods Pond (Figure 6.2-50) as suggested by the overprediction of the majority of 1999 data. The failure of the model to reproduce the rate of decline in sediment PCB concentrations is a major problem, since in the long-term, PCB concentrations in biota will respond to this rate of change. The MVR fails to address this discrepancy and it's possible explanations, even though this error is evident in both calibration and validation periods. It would also be useful to compare sediment PCB concentration predictions to data for deeper, sub-surficial layers. Slower changes in PCB concentrations would be expected in deeper intervals, reflecting slower transport in deeper sediments.

Response 2A-DE-2:

As discussed above in the response to this General Topic, EPA conducted further statistical analyses of potential temporal/spatial trends in water column, sediment, and fish tissue PCB concentrations; these analyses are reported in detail in Appendix A.1 of the FMD. Based on these more in-depth statistical analyses, which incorporate a correction for known bias in some of the PCB laboratory analyses, the decline in measured sediment PCB concentrations throughout Reaches 5 and 6 during the 26-year validation period was shown to be slow and not significantly different from zero. The model simulation is consistent with the results of these analyses.

A number of significant model biases have been identified and discussed above (Question 2). The most significant are:

- The predicted changes in sediment PCB concentrations are slower than observed,

Response 2A-DE-3:

The statistical analysis of sediment PCB concentrations (Appendix A.1 of the FMD) indicates that there is no statistically significant change in sediment concentrations with time. The rate of sediment PCB concentration decline predicted by EFDC is slow and is well within the uncertainty envelope of the data. The Reviewer's interpretation of an apparent trend in the data, leading to a conclusion of a significant bias in the model, is incorrect and inconsistent with the statistical analysis results and the interpretation of the data presented in the MVR (Section 6.2.3.4).

- Spatial gradients in water column PCB concentrations do not match data under low flow conditions, and

Response 2A-DE-4:

PCB mass transfer between the bed and water column is proportional to the PCB concentration gradient between the water column and sediment pore water. This concentration gradient does not substantially change in response to the upstream PCB boundary condition because pore water PCB concentrations (on the order of 10 µg/L) are much larger than water column PCB concentrations (approximately 0.05 to 0.10 µg/L). This means that the gradient (the pore water value minus the water column value) is approximately equal to the pore water PCB concentration. As a result, the PCB mass transfer from sediment is not noticeably affected by the upstream boundary concentration.

The calculated incremental PCB concentration increase between the confluence and New Lenox Road (over a flow range from 50 to 100 cfs) that results from the sediment PCB mass transfer is comparable in magnitude to the concentration at the upstream boundary during low-flow periods. Consequently, the uncertainty associated with setting the boundary concentration does not markedly alter the ability of the model to simulate the water-column PCB concentrations. For example, if the upstream PCB boundary concentration were set to 0.01 µg/L, the water column PCB concentration would be controlled by the PCB mass transfer from the sediment and would produce a water column PCB concentration of approximately 0.10 µg/L at the downstream end of Reach 5A. Alternatively, if the boundary concentration were set to 0.10 µg/L, the water column PCB concentration at the end of Reach 5A would be approximately 0.2 µg/L. As a result, the sensitivity of the mass transfer flux to changes in the upstream boundary condition is small.

- The treatment of the upstream boundary condition results in serious overprediction of low TSS and PCB concentrations.

Response 2A-DE-5:

EPA disagrees with the Reviewer's assessment that the treatment of the upstream boundary condition results in a serious overprediction of low TSS and PCB concentrations for the reasons stated above.

Marcelo H. Garcia

Of particular relevance for this reviewer is the validation of flood plain deposition throughout the river system as well as sediment erosion and deposition in Woods Pond, since the record shows that this pond is a major deposit for PCB-contaminated sediments.

Regarding the sedimentation of Woods Pond, the calibration/validation exercise has produced results showing rates of (*net*) sedimentation for the period going from 1979 to 2004 (Figure 6.2-34). The results seem reasonable with predicted sedimentation rates (mm/year) that are within an order of magnitude of sedimentation rates determined from Cesium data. The model also reproduces fairly well the distribution of PCBs with depth as observed from sediment cores. However, as stated below, the model shows some bias when predicting temporal variations of surface PCB concentrations in Woods Pond.

Woods Pond is perhaps one of the few locations in the Housatonic River where the capabilities of the model to predict the spatial and temporal distribution of sediment and PCB can be extensively tested due to the large amount of observations available. This has been done to some extent but more effort should go into this since the pond could become a major source of PCBs to the downstream portion of the river during a major flood event.

Response 2A-MG-1:

Additional comparisons between simulated and measured sedimentation rates in Woods Pond, which were not included in the MVR, have been included in the FMD (Section 4 and Appendix B.6). A more detailed statistical analysis of temporal distributions of PCBs in Woods Pond sediment has been completed, and is presented in Appendix A.1 of the FMD. EPA agrees with the Reviewer's comment that the comparisons between measured and simulated sedimentation rates are reasonable, and points out that the agreement on a core-by-core comparison is better than the order-of-magnitude characterization mentioned by the Reviewer.

As shown in MVR Figure 6.2-50, the model predictions of variation with of surface PCBs concentrations are not consistent with the field observations in Woods Pond. The model seems to underpredict the observed values in the time period from 1979-1984 and to overpredict the observations in the time period from 1995 to 2000.

Response 2A-MG-2:

As stated in response to comments from other Reviewers, a statistical analysis of Woods Pond sediment PCB data, described in Appendix A.1 of the FMD, indicates that there is no statistically significant change in sediment concentrations with time, and that the slow simulated rate of sediment PCB concentration decline is well within the uncertainty envelope of the data.

There is evidence that model predictions do not match well with observed surface sediment concentrations of PCBs in reaches 5A, 5B, and 5C as displayed in Figure 6.2-49.

Response 2A-MG-3:

EPA disagrees with the Reviewer's comment that the model predictions are not consistent with the data. The statistical analysis described in Appendix A.1 of the FMD and previous characterizations of the trends (or lack thereof) in PCB sediment concentrations indicate there are no trends in the data, and that there is a great deal of variability in the data. The model simulates the central tendencies of the data very well.

Frank Gobas

The model calibration has been complicated by the lack of a significant change in PCB concentration data in the River during the calibration period. Hence, the model calibration sheds little light on the issue of whether the model is able to account for changes in PCB concentrations over time that may occur as a result of remediation. The main conclusion from the calibration is that the temporal response of PCB concentrations in the River is likely slow, but it cannot tell us accurately how slow.

Response 2A-FG-1:

EPA agrees that the data do not suggest a significant change in PCB concentrations, and this is supported by the conclusions of a statistical analysis of the data provided in Appendix A.1 of the FMD as well as previous characterizations of the trends (or lack thereof) in PCB concentrations in sediment. EPA agrees with the conclusion that the Reviewer drew from the Calibration that the temporal response of the system is slow, and acknowledges that there is uncertainty in quantifying how slow the rate is.

Evidence to indicate that the current temporal response of the model may be flawed was presented by Dr. Connolly, who suggested that the half-life time of PCBs in the River is likely is closer to 36.5 yr (i.e. 0.693/0.019) than the current model prediction of 110 yr (i.e. 0.693/0.006). I think that Dr. Connolly is probably correct, but the uncertainty in the estimates of the half-lives is large and I have not been convinced that the current PCB concentration data in the River allow one to distinguish between these two estimates.

Response 2A-FG-2:

While there is uncertainty in the estimates of the rate of decline in sediment PCB concentrations, EPA disagrees with the Reviewer's opinion that Dr. Connolly's estimate of a half-life of 36 years is probably correct, rather than the simulation results of 110 years. As stated previously, a statistical analysis of Woods Pond sediment PCB data (described in FMD Appendix A.1) indicates that there is no statistically significant change in surficial sediment PCB concentrations with time, and that the slow simulated rate of sediment PCB concentration decline is well within the uncertainty envelope of the data.

Because the sediment PCB concentrations in the Housatonic River exhibit such a high degree of spatial variability, the number of data points available to calculate an annual average is important. The number of data points available to calculate each annual mean is included on the temporal plots of sediment concentrations (FMD Figure 3.5-20) to provide the reader with an indication of the number of samples used to estimate an annual average concentration in a particular reach of the river. The analysis presented by Dr. Connolly did not consider confidence bounds (which are a function of the number of samples and variability of the concentrations) around the annual mean concentrations used to assess trends over time. When the number of samples and the wide confidence bounds are considered, it is not surprising that the conclusions from the rigorous statistical analysis contradict those presented by Dr. Connolly. In fairness to Dr. Connolly, he characterized his analysis as "dumb and simple."

In conclusion, the current characterization of the key temporal aspects of the model is too uncertain to make reliable long term predictions of the concentrations of PCB in water, sediments and biota of the Housatonic River. This said, it is clear from the empirical data and the model calculations that the long term temporal response to changes in PCB loadings is slow. There appears to be very large standing mass of PCBs in the River sediments and its floodplains and the sediment removal rate from the PSA is relatively small. Natural recovery therefore can be expected to take a long time, with a system half-life time in the order of decades. Given this slow response time, the calculation of a more accurate natural recovery time may in some cases be inconsequential. Under certain conditions, the model, as it stands, may already be sufficient to address the long term temporal response of PCB concentrations in the water column and bed sediments (model objective #1).

The reason for the model's ability to convincingly assess the long term temporal response of the PCB concentrations in water, sediments and biota of the River are twofold. First, no significant long-term-changes in PCB concentrations over time were observed in water, sediment or biota throughout the PSA during the calibration period. Secondly, the key parameters controlling the temporal response of PCB concentrations are difficult to measure or estimate.

Response 2A-FG-3:

EPA agrees with the Reviewer's characterization that it is clear from the data and the model results that the long-term temporal response to changes in PCB loadings is slow, and that there is a large inventory of PCBs in the sediment and its floodplain, and that the rate of sediment decline in the PSA is relatively small. Statistical analyses and model simulations suggest that the system half-life time is on the order of decades, as characterized by the Reviewer. EPA acknowledges that given the uncertainty in the rate of decline (due to the lack of ability to project this from the data), relative predictions by the model are likely more reliable than absolute predictions. This is reflected in the statement of the goal of the modeling study.

Although factors affecting the temporal response of the system are difficult to measure or estimate, considerable effort has gone into parameterization of these factors using site-specific data and information from literature (e.g., benthic biota, coupled with literature values and SPI work for rates and depths of mixing; sedimentation rates for net burial).

The modelers should revisit the calibration of the temporal response of the model based on a more in-depth statistical analysis of the available model calibration data. I specifically refer to the sediment concentration data as a function of time. Model calibration schemes should be evaluated in terms of their ability to reproduce statistically validated temporal PCB concentration data.

Response 2A-FG-4:

EPA agrees with the Reviewer's suggestion. The recommended statistical analyses have been conducted (see Appendix A.1 to the FMD), and these analyses indicated that the simulated rate of change in sediment PCB concentrations by the model are consistent with the those observed in the data for both sediment and fish tissue PCB concentrations.

The report indicates that the model reasonably accounts for the role of storm events. The model appears to do a good job in describing the short term temporal response of river flow rates in response to storm events. The comparison of simulated and observed river flows is more than adequate. However, the excellent flow predictions do not translate in equally impressive TSS concentrations. Figures 6.2-19 to 6.2-22 show considerable discrepancies between observed and predicted short term TSS concentrations at Holmes Road, Lennox Road and Woods Pond Outlet. On the other hand, Figures 6.2-39 to 6.2-41 show that the short term PCB concentration variations in the water phase (assuming that the concentration data displayed are water column tPCB concentrations) are reasonably well predicted. The reason for the poor predictability of TSS concentrations, but good predictability of the PCB concentrations in the water column is somewhat surprising and unclear. However, the data presented produce confidence in the model's ability to make reasonable predictions of the effects of storm event(s) on the redistribution of PCB-laden sediment in the study area (model objective #4).

Recommendation:

I recommend that the modelers investigate the source of error in the estimation of the TSS.

Response 2A-FG-5:

EPA agrees with the Reviewer's comment that the model's ability to make reasonable predictions of the effects of storms has been demonstrated. The modeling team has thoroughly investigated the source of the differences between simulated and measured TSS, which are related to two major factors.

The first factor is that the model is designed to represent the typical (average) response of the river to changes in flow and loadings. Because it represents the average response, simulated TSS concentrations are less variable than those measured. At low measured TSS concentrations (2 mg/L to 4 mg/L), simulated concentrations are typically higher than those measured, but generally within 2 to 4 mg/L of the data. Closer agreement between simulation and measured concentrations is generally achieved for higher TSS concentrations, which typically occur at high flows. These periods are important for both solids and PCB transport because these conditions result in the largest loadings. Overprediction of low TSS concentrations that occur at low flow are less important because loads for these conditions are far smaller (although on the logarithmic scale, may appear large). Further, even though TSS concentrations at low flow could also impact PCB concentrations at low flow, exposure concentrations estimated by EFDC are not strongly influenced by differences between simulated and measured TSS concentrations at low flow. As part of the evaluation of the significance of the differences between simulated and measured concentrations at the low range of measured concentrations, a bounding analysis was performed using FCM, as discussed in further detail in Section 4.3 of the FMD. The conclusions from the bounding analysis are that the overpredictions of the low TSS and PCB concentrations have small effects on solids mass transport, exposure concentrations, and resulting fish tissue PCB concentrations.

The second factor that contributes to differences between simulated and measured TSS concentrations is the effect of concentrations assigned to upstream boundary conditions on the model simulations when data are not available. This factor is most significant for Holmes Road, which is close to the model upstream boundary. Agreement between simulated and measured TSS is considerably better at New Lenox Road than at Holmes Road, especially in the Phase 2 calibration period when data for New Lenox Road cover a wider range of flow conditions and boundary condition estimation effects are less important. The modeling team devoted considerable effort to developing methods to estimate upstream boundary conditions for times when data were not available (a full discussion of these efforts is provided in Appendix B.2 of the FMD). Large variability in measured boundary condition concentrations at low flows resulted in both positive and negative differences between measured and estimated boundary concentrations, because the estimates from the regression equations reproduce the central tendency of the data.

Based on the results of the bounding calculations, the differences between simulated and measured TSS concentrations were judged to be acceptable because of the small effect on TSS and PCB mass fluxes.

Thirdly, the analysis presented by Dr. Connolly during the June 28 Public Meeting indicates that there may be a bias in the temporal response of the model, i.e. it appears that the rate of temporal response of PCB concentrations in the River is underestimated by the model. However, there is considerable uncertainty in the data on which the calculations are based and it is unclear whether the uncertainty in the data is sufficiently small to distinguish between the temporal response rates calculated by the model and Dr. Connolly's analysis. A bias in the temporal response of the model could have large implications for model projections of remedial actions. More detailed evaluations of the bias in temporal response need to be conducted. This should involve the calculation of PCB concentration response times from observed data and the comparison of the measured response times to the calculated response times to determine whether a bias exists. Alternative model parameterization schemes may need to be explored to investigate whether the model bias (if it exists) can be reduced.

Response 2A-FG-6:

EPA disagrees that there is a bias in the temporal response of the model for the reasons stated in the responses above. The Reviewer is correct in recognizing the effect of uncertainty in the data on estimation of temporal response. The statistical analysis recommended by the Reviewer had been conducted, and confirms that there is no temporal bias in model simulations.

The possible bias in the model's temporal response of the PCB concentrations in the River needs to be investigated by comparing observed and predicted PCB concentration decline rates over time. In case, significant bias exists, alternative model parameterization schemes need to be explored to improve the long term temporal response of the model.

Response 2A-FG-7:

The results of the statistical analysis of surface sediment PCB data supports the temporal response simulated by the model and indicates no temporal bias.

With regard to the temporal scale, the comparisons of model predictions with data are not sufficient to evaluate the capability of the model to make accurate estimates of the temporal response of PCB concentrations in the River. As discussed under charge question #1, this is largely due to the considerable variability in observed PCB concentrations and the lack of a significant decline in PCB concentrations over the calibration period. Hence, a temporal trend is difficult to discern from the data and the lack of temporal trends makes it very difficult to evaluate the temporal characteristics of the model.

Response 2A-FG-8:

The Reviewer is correct in commenting that the variability in observed PCB concentrations and lack of a significant decline in concentrations make evaluation of the temporal characteristics difficult. It is noted, however, that the model results are consistent with both the slow change in sediment concentrations (confirmed by statistical analysis – FMD Appendix A.1) and the sedimentation rates derived from an analysis of cesium-137 data.

With regards to the spatial scale of the model, a similar conclusion can be reached at. The small-scale spatial variability in the PCB concentration in the sediments is so great that spatial differences in PCB concentrations among River sections are difficult to discern. The comparison between observed and predicted concentrations therefore provides little information with regards to the capability of the model to make accurate predictions of PCB concentrations as a function of time or space.

Recommendations:

Dr. Connolly's analysis suggests that with the application of suitable statistics it may be possible to use the current data sets to better characterize temporal and spatial PCB concentration trends. Spatial statistics and the application of geographical averaging methods may help to better characterize spatial trends in the data that can be used to evaluate the applicability of the model. Temporal trend analysis can be used to discern temporal trends. I recommend that this is done as it may provide better data for a comparison of observed and predicted concentrations.

Response 2A-FG-9:

The statistical analyses recommended by the Reviewer have been conducted (see FMD Appendix A.1) and support the simulated rate of change in surface sediment PCB concentrations.

The model is the only available tool to simulate the future response of PCB concentrations in response to remediation efforts. The model framework represents a suitable approach to estimating the future time response of PCB concentrations in the River and the calibration and validation of the model have involved significant efforts.

Response 2A-FG-10:

EPA agrees with the Reviewer's comment.

The most valuable information for the calibration and validation of the model is a change in PCB concentrations in the River in response to a known reduction in PCB loading. This kind of information was not obtained in the current study as concentrations of PCBs in the River showed little or no significant variation with time. The current model therefore has to rely on the characterization of a number of key state variables for the estimation of the long term temporal response of PCB concentrations to remedial scenarios. The key parameters include the amount of "available" River and floodplains sediments and rates of resuspension, diffusion, bioturbation and subduction. All of these model state variables are either currently unmeasurable or very difficult to measure or estimate. As a result the model's outcome with regards to the long term time response of PCB concentrations in the River is uncertain. The model uncertainty translates in considerable uncertainty about future PCB concentrations in the River resulting from remediation efforts or the absence of remediation.

Response 2A-FG-11:

EPA acknowledges that there will be uncertainty associated with future model predictions. In any model framework, uncertainty arises from (1) the data and (2) the degree of process aggregation (model formulations). Regardless of the degree of process aggregation, no model can have less uncertainty with its predictions than the uncertainty of the data.

However, it should be recognized that the response of the model to changes in PCB loadings is quantified. As was shown in the RFI, other documents, and is more fully explored in Appendix A.1 to the FMD, the rate of change in surface sediment PCB concentrations in Reaches 5 and 6 (Woods Pond) is not statistically different from zero (i.e., slow to no change over time). The example and tracer simulations were performed explicitly to demonstrate the model response.

Wilbert Lick

A good comparison of the model predictions with data is necessary, but not sufficient, to evaluate the capability of the model. The low surficial PCB concentrations in Woods Pond and the large variability in PCB concentrations throughout the river are unexplained by the present model. See discussion in the complete response.

Response 2A-WL-1:

It is assumed that the Reviewer's statement that "low surficial PCB Concentrations in Woods Pond … are unexplained by the present model" refers to the temporal comparison of simulated and measured concentrations. As stated in responses above, the statistical analysis of Woods Pond surface sediment PCB concentrations, described in Appendix A.1 of the FMD, indicates that there is no statistically significant change in sediment concentrations with time and that the rate of PCB concentration decline simulated by the model is slow and is well within the uncertainty envelope of the data.

With respect to the Reviewer's comment regarding unexplained large variability in PCB concentrations throughout the river, it is noted that this variability was observed on a spatial scale of 1 m^2. Because of this variability, a decision was made to develop average initial condition for spatial bins of approximately 400 meters in length (0.25 miles). Average concentrations were calculated for each spatial bin and assigned to those grid cells (approximately 20 meters in length) within the bin. Assigning a single average concentration to all grid cells in a spatial bin, by definition, reduces the variance of the PCB concentrations represented in the initial conditions. Although cell-to-cell variations in PCB concentrations developed in the model over time due to cell-to-cell variations in erosion and deposition (MVR Figure 4.2-79), it is completely expected that the simulated concentrations have less variability than the original data used to calculate the average initial conditions. This approach was adopted because the goal of the modeling study is addressed at spatial scales greater than 1 m^2, or 20-meter grid cells.

During previous peer review meetings, John List as well as others including myself (but John was most vocal) have emphasized (a) the large unexplained variance in the PCB concentrations in the surficial (six inch) layer of the sediments and (b) the unexplained high concentrations of PCBs in the surficial layers of the sediments in Woods Pond.

EPA had no explanations for these latter two problems, but also stated that an understanding of these problems was not necessary. An understanding of these problems may not be necessary, depending on your point of view, but the problems themselves are quite interesting, deserve some discussion, and are related to the inaccurate modeling of the basic sediment and contaminant flux processes mentioned above. Some discussion of these problems is given here.

Response 2A-WL-2:

EPA agrees that the issue of the variability in sediment PCB concentrations is interesting and it has been discussed in each of the reports developed during the modeling study (MFD, MCR, MVR) and each of the modeling Peer Reviews. The EPA modeling team disagrees with the Reviewer's comment that the unexplained variance in surface sediment PCB concentrations is related to inaccurate modeling because the same variability is noted on a very small-scale in the benthic sampling data set, in which variability of two orders of magnitude or more was observed at each of 9 sampling locations where 12 samples were collected within an area of 1 m^2. Attempting to model this system with a grid resolution on the order of 0.1 m^2 (1 m^2/9 samples) clearly is not feasible or necessary to achieve the goal of the modeling study.

The higher than predicted PCB concentrations in Woods Pond as well as much of the variability seen in the sediment PCB concentrations can be explained by finite PCB sorption rates as well as by the variability in the PCB sources and the hydrodynamics. The improved model should be able to predict some of this and at least suggest the reasons for the remaining variability if finite sorption rates are assumed and a finer grid is adopted.

Response 2A-WL-3:

EPA disagrees with the Reviewer's comment. First, as discussed in previous responses, the PCB concentrations simulated are not higher than those "predicted" by the data. Second, finite PCB desorption rates would result in minimal (<10%) change in concentrations of PCBs on deposited solids.

E. John List

A similar point arises in connection with the Validation Report, on Page 4-12 from line 23 et seq. through page 4-13. The high concentrations of PCB in the surface layers of the sediment in Woods Pond (apparent also in Figure 4.2-79) and the lack of an explanation thereof are troubling. Again EPA has thrown up its hands and states that no explanation has been found (see p. 4-13 of the Validation Report).

Response 2A-JL-1:

EPA disagrees with the characterization that "EPA has thrown up its hands" with respect to evaluating the small-scale variability. Rather, EPA has determined that explaining the source of the small-scale variability is not necessary for achieving the goal of the modeling study, and is not worthy of a research effort beyond the scope of the modeling study.

I recognize the fact that these a very difficult problems to address, but if the modeling is to provide believable predictions of the future fate of the PCB in the sediments it seems to me that the processes that have resulted in the existing PCB concentration distributions have to be understood first.

Response 2A-JL-2:

See Response 2A-JL-1 above.

With respect to the PCB modeling the results are so tenuous (see Figure 6.2-49) that it is not possible to draw any really solid conclusion with respect to bias. However, an important point is demonstrated by Figure 6.2-50. This figure indicates that the sediment concentrations of PCB in the lower reaches of the project study area have declined somewhat since 1990 (i.e., since remediation activity started). Furthermore, there appears to be a similar decline in the water column concentrations of PCB, as is somewhat evident in the lower panel of Figures 6.2-44. (It is noted that there is no statistical measure of the significance of the decline in either the sediment or the water column, but the data are very suggestive). However, the modeling does not give any indication that any similar rate of decline has occurred (see Figure 6.2-50). In fact the comparison between the data and the model in this figure is strongly suggestive that the model does have a bias.

Response 2A-JL-3:

The Reviewer states that the data shown on Figures 6.2-49 and 6.2-50 are suggestive of statistically significant declines in sediment PCB concentrations. The data are presented as the mean ± 2 standard errors, which approximates the 95% confidence interval of the mean. At that time, the wide range in the confidence interval of the mean and the small number of samples in earlier years were identified by the modeling team as factors that complicated characterization of trends in the data and comparisons with trends computed by the model.

> In response to the misconception of the Reviewers regarding the existence of a trend in PCB concentrations in sediment, a more rigorous statistical analysis was performed (see Appendix A.1 of the FMD). This analysis confirms previous conclusions that there is no statistically significant decline in sediment PCB concentrations. The results from this (and prior) analyses contradicts the basis for the Reviewer's conclusion that these model-data comparisons indicate a bias in the model.
>
> In addition, it should be noted that the Reviewer is incorrect in his observation that remediation began in 1990; no remediation was conducted in the river until 1997 (the Building 68 Removal Action), and significant continuous remediation did not begin until October 1999.

The bias of the model is further reinforced by the results of the hypothetical remedial scenarios plotted in the un-numbered figure on page 56 of the EPA Response to Questions from Model Validation (Document DCN: GE-061406-ADFI), where, despite a reduction in the flux of PCB from the sediments in the upper river reaches, there is no significant change in projected sediment concentration of PCBs in the lower reaches. Thus the model does not match the reduction in sediment concentration that has occurred subsequent to actual post-upstream remediation, and projects no reduction as a result of an hypothetical reduction in the upstream PCB flux, as in the hypothetical examples. This problem needs to be understood and/or fixed before the model is used in real applications.

Response 2A-JL-4:

> The results from the simulations of the example scenarios show a gradual decline in PCB concentrations in sediment in downstream reaches for the base case and Examples 1 and 2 (which include elimination of upstream sources). The differences among the rates of decline in sediment PCB concentrations in the base case and example simulations are small, which is consistent with the statistical analysis that concludes that conditions in the river are changing slowly.
>
> It must be noted that the Reviewer's statement that impacts of remediation should have an appreciable influence on the model results is incorrect. With the exception of limited activity associated with Building 68, the remediation in the ½-Mile Reach began in October 1999, and remediation in the 1½-Mile Reach began in 2002. While the model simulation used the 26-year validation hydrograph, the initial conditions were derived using data that pre-dated remediation. The Reviewer's evaluation of example scenario simulations is based on the incorrect assumption that the data reflect 9 years of remediation activities.

2.2 GENERAL TOPIC 2B: SMALL-SCALE SPATIAL VARIABILITY

2.2.1 Comment Summary

Reviewer Name	Page Number/Line Number
Adams	Page 5, lines 14 to 37
Endicott	Page 4, lines 11 to 20
	Page 8, lines 13 to 21
Garcia	Page 3, lines 8 to 13
Gobas	Page 4, lines 10 to 15
	Page 4, lines 35 to 36
	Page 10, lines 27 to 32
Lick	Page 1, lines 30 to 33
	Page 10, line 38 to page 11, line 3
List	Page 3, lines 10 to 21
	Page 3, line 22 to page 4, line 3
	Page 4, lines 18 to 21

2.2.2 General Topic Summary

A high degree of small-scale spatial variability in PCB concentrations in surface sediment was observed in data collected in the benthic sampling program, in which variability in concentrations of two orders of magnitude or more was observed at each of 9 sampling locations, where 12 samples were collected within an area of 1 m^2. Similar variability is observed in the remainder of the surface sediment data set, and is discussed in Appendix A.1 to the FMD and in previous model documents including the MFD, MCR, and MVR.

Six of the seven Reviewers commented on the variability in PCB concentrations over small spatial scales, particularly in regard to sediment PCB concentrations. In some cases, as noted in the summary of General Topic 2A, these comments were made in the context of the difficulty of understanding spatial patterns and temporal trends in the presence of such underlying variability.

Reviewers were divided on the implications of the acknowledged inability of the model to reproduce the observed small-scale variability in sediment PCB concentrations on achieving the goal of the modeling study. One Reviewer did not feel that the lack of such variability in model predictions was an issue as long as it was acknowledged. Two Reviewers indicated that the issue was of major importance and an indication that the model does not adequately reproduce the underlying processes. The remaining Reviewers indicated the lack of small-scale variability in model output to be problematical, but not sufficient to indicate major flaws in the underlying structure or use of the model.

2.2.3 General Topic Response

EPA has spent considerable effort attempting to understand processes responsible for the small-scale variability. Even though no mechanistic explanation has been developed, EPA does not agree with the Reviewer's comment that the lack of such an explanation compromises the ability of the model to achieve the goal 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. In addition, the sampling programs were not designed to assess short-term temporal variability and any such short-term temporal variability that does exist would be masked by the spatial variability.

EPA believes that it is not necessary to understand the processes that underlie the small-scale variability or to include such processes or the resultant variability in the framework to achieve the goal of the modeling study. As described in previous documents and in the FMD, the spatial scale of interest is larger than a grid cell and smaller than a reach. 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, and it is not necessary to do so. 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.

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 m^2. Results of PCB analyses of 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 coating, which was identified by the SNL 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 a consequent reduction in bioavailability. This was addressed in FCM for Reach 5A. This was analyzed and discussed at length in the Model Calibration Report (Section C.3.1.2) and a procedure was developed to arrive at adequate calibration.

2.2.4 Responses to Specific Comments on General Topic 2B

Eric Adams

Much has been 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. Whether or not this is a model failure, *per se*, or simply unresolved variability in model input and output (sediment bed concentrations), depends on what has caused the variability. I believe the variability was caused mainly by the stochastic method in which the PCBs were introduced in the first place. In such case, 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 (and some non-linearity is inevitable, given the averaging associated with the coarse grid).

Response 2B-EA-1:

EPA agrees with the Reviewer's comment that the model cannot be expected to predict the variability observed in the data. In addition, 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.

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 (at least for natural attenuation).

Response 2B-EA-2:

EPA agrees that the variability in the existing data could provide an estimation of the variability in future concentrations (with some important caveats and the recognition of significant uncertainty). It should be noted that the timeframe of interest will likely be longer than that presumed by the Reviewer.

On the other hand, if the variability is due to active sediment transport processes that are sorting the sediments and their contaminants, then the failure to pick this up could be a significant model deficiency. For example, natural attenuation could conceivably increase local PCB concentrations. The available time series data of PCB concentrations within surficial sediments suggest a decrease in concentration (indeed more so than is being modeled), so I don’t believe this is a significant process.

Response 2B-EA-3:

EPA agrees with the Reviewer that analysis of the data shows no statistically significant trend and that the model simulates a very slow decline; therefore, it is not a significant process.

Douglas Endicott

EFDC uses the same equilibrium partitioning (EP) model as every other fate and transport model in use today. EP generally works quite well, and simplifies the model computation by requiring only a single state variable for each chemical. It has been suggested that desorption kinetics should be incorporated into the Housatonic River model, because PCBs desorbing from resuspended sediment might not reach equilibrium within the timeframe they remain in suspension. This disequilibrium may help explain the pattern of PCBs moving in “ribbons” of sediment downstream, and the high surficial PCB concentrations in Woods Pond sediment. I think this is mostly speculation and, although I am curious to see if it is true, it is well and beyond the objectives and goals of this modeling study.

Response 2B-DE-1:

EPA agrees with the Reviewer that non-equilibrium partitioning is beyond the objectives and goals of the modeling study, and has presented calculations in FMD Appendix B.9 to support the appropriateness of the use of equilibrium partitioning. It is noted that there is no evidence of a pattern of PCBs moving in “ribbons,” but rather the data suggest that concentration patterns are random.

Figures 4.2-79 and 6.2-51 are intriguing plots of the change in spatial profiles of the grid cell PCB sediment concentrations over the duration of the calibration and validation, respectively, superimposed with the “shotgun pattern” of data for individual sediment PCB samples. Although no interpretation of the model-to-data comparison is offered, it does suggest that the greater than 90% change in sediment PCB concentrations within many sediment grid cell is both plausible and may help explain the spatial variability observed in sediment PCB data throughout the river sediments. In other words, the model simulation suggests that sediment PCB variability arises from contaminated sediments remaining in locations where erosion is not significant, and (relatively) uncontaminated sediments replacing older contaminated sediments in locations where erosion is significant.

Response 2B-DE-2:

The Reviewer correctly notes that spatial variability in simulated PCB concentrations increases through the course of the simulation in response to variability in erosion and deposition that is caused by local bathymetric features. It is noted, however, that there is almost as much variability in the samples collected within 1 m^2 where no variation in erosion and/or deposition patterns are expected, as shown in MVR Figures 4.2-79 and 6.2-51. Therefore, EPA does not believe that erosion and deposition patterns alone explain the small-scale variability in PCB concentrations.

Marcelo H. Garcia

While the model capabilities have been improved since its calibration, my professional opinion is that the validation of the model is not complete at this stage. In particular the model does not capture the variability and spectrum of PCBs concentrations observed in river. There are several aspects of the model that need to be improved as stated above before validation of the model can be accomplished.

Response 2B-MG-1:

EPA 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.

Frank Gobas

The model report reveals insufficient information to determine whether the capability of the model to make accurate predictions of spatial differences in PCB concentrations is adequate. The reason is that the calibration of the spatial characteristics of the model revealed that differences in PCB concentrations (e.g. see Fig 6-2-45 and 6-2-46 Model Validation Report) among River sections are small while variability in the PCB concentrations in the sediments at individual river locations is high.

Response 2B-FG-1:

The Reviewer is correct in noting that the variability in the data complicates evaluation of spatial and temporal trends. The figures identified by the Reviewer in his comments regarding spatial patterns present model-data comparisons for low-flow conditions (flows at Coltsville between 16 and 25 cfs) and conditions slightly above average flow conditions (flows at Coltsville between 109 and 275 cfs). These figures highlight the effect of the spatial transient in PCB concentrations due to PCB mass transfer from the sediment under low- to moderate-flow conditions. An additional and perhaps more important demonstration of model performance is the ability of the model to reproduce the roughly one order of magnitude decline in water-column PCB concentrations and also the decline is TSS concentrations between New Lenox Road and Woods Pond Outlet under high-flow conditions (Figures 3.5-6 and 3.5-15 in the FMD, previously presented as Figures 4.2-75 and 4.2-76 in the MVR).

Better spatial statistical methods should be used to explore the current spatial PCB concentration data for spatial trends in the available PCB concentration data.

Response 2B-FG-2:

This analysis has been performed and is presented in Appendix A.1 of the FMD.

I agree with the EPA that it is not necessary to understand the processes causing small scale variability in concentrations (l.3-4, p.15 Model Validation Report). This is normal is any scientific observation. However, when interpreting the observation or the model calculation (in this case), it is then important to recognize the uncertainty that is associated with the lack of understanding, so that it can be taken into account in the decision analysis.

Response 2B-FG-3:

EPA agrees with the Reviewer's comment and the uncertainties associated with the small-scale variability and the effect on the model simulations are recognized. See FMD Sections 3 and 4, and Appendix A.1.

Wilbert Lick

A good comparison of the model predictions with data is necessary, **but not sufficient**, to evaluate the capability of the model. The low surficial PCB concentrations in Woods Pond and the large variability in PCB concentrations throughout the river are unexplained by the present model. See discussion in the complete response.

> **Response 2B-WL-1:**
>
> As indicated in the General Response, EPA believes that it is not necessary to explain the large variability in PCB concentrations to achieve the goal of the modeling study.

During previous peer review meetings, John List as well as others including myself (but John was most vocal) have emphasized (a) the large unexplained variance in the PCB concentrations in the surficial (six inch) layer of the sediments and (b) the unexplained high concentrations of PCBs in the surficial layers of the sediments in Woods Pond.

EPA had no explanations for these latter two problems, but also stated that an understanding of these problems was not necessary. An understanding of these problems may not be necessary, depending on your point of view, but the problems themselves are quite interesting, deserve some discussion, and are related to the inaccurate modeling of the basic sediment and contaminant flux processes mentioned above. Some discussion of these problems is given here.

> **Response 2B-WL-2:**
>
> EPA agrees that the issue of the variability in sediment PCB concentrations is interesting, and it has been discussed in each of the reports developed during the modeling study (MFD, MCR, MVR) and each of the modeling Peer Reviews. The modeling team disagrees with the Reviewer's comment that the unexplained variance in surface sediment PCB concentrations is related to inaccurate modeling because the same variability is noted on a very small scale in the benthic sampling data set, in which variability of two orders of magnitude or more was observed at each of 9 sampling locations, where 12 samples were collected within an area of 1 m^2. Attempting to model this system with a grid resolution on the order of 0.1 m^2 ($1m^2/9$ samples) clearly is not feasible or necessary to achieve the goal of the modeling study.

E. John List

It is my professional opinion that the model does not "reasonably account for the relevant processes", nor, *in its present state*, can it reliably "quantify future spatial and temporal distributions of PCBs".

> **Response 2B-JL-1:**
>
> Many of the Reviewer's comments focus on the need to explain the small-scale variability of sediment PCB concentrations and an erroneous interpretation that the sediment data describe a statistically significant decline in sediment concentrations. EPA has repeatedly stated that explaining the cause of the small-scale variability is not necessary to achieve the goal of the modeling study. In addition, EPA has also noted that there is no statistically significant trend in the surface sediment data in the MFD, MCR, and MVR, and this was discussed in the RFI. EPA has also conducted further rigorous statistical analysis that

> definitely shows no statistically significant change in PCB concentrations in surface sediment.

One of the reasons for this opinion is encompassed in Figure 4.2-79 of the validation report. This figure indicates, from measured data, that 0-6 inch sediment PCB concentrations within the river and flood plain are spread over three orders of magnitude (0.2-200 mg/kg) at locations that are very close together along the river. Furthermore, for almost 12 miles along the river the variance in the distribution of concentrations appears to be fairly uniform. From the fact that the PCB concentrations appear this way after 30-40 years of river action involving sediment erosion and deposition it is reasonable to conclude that the sediment transport processes are not smoothing the sediment PCB concentration distributions.

> **Response 2B-JL-2:**
>
> EPA agrees that the sediment transport processes are not smoothing the sediment PCB concentration distributions. In fact, Figure 4.2-79 of MVR demonstrates that through the course of the model simulations, variability in sediment concentrations increase relative to the initial conditions, which were established as averages within spatial bins. The increase in spatial variability is due to variability in erosion and deposition caused by spatial variations in bathymetry.

A second reason for this opinion is that recent river bottom survey profiles indicate quite clearly that sediment erosion and deposition is very definitely non-uniform across the river cross-section. By contrast, and as discussed above, the modeling uses a single computational cell to define the entire width of the normal river channel and the local erosion is defined by the fluid shear stress that is averaged across the river cross-section. Because of the non-uniform nature of the actual erosion, and the fact that the model uses an average shear stress to define the erosion, it is highly probable that river bed erosion occurs long before the model predicts its occurrence. Furthermore, the PCB concentration of the sediment that is eroded in the model is the average of the local concentration over that cell. Thus the modeling uses cell averages that reduce the very broad PCB concentration distributions to a local average, and then redeposits that PCB-laden sediment at this average concentration. The net effect of the modeling therefore must be to smooth out the distribution of PCBs, but in reality this has not occurred in the significant length of time in which, according to the model, it should have occurred.

> **Response 2B-JL-3:**
>
> EPA disagrees with the Reviewer’s assertion that cross-channel patterns in erosion and deposition in the Housatonic River are always nonuniform. Although there are exceptions, cross-channel differences in erosion and deposition are frequently minimal (21 of 53 profiles examined), as described in Appendix B.6 of the FMD. The Reviewer’s conclusion that the net effect of modeling smoothes out the distribution of PCBs is incorrect. In fact, through the course of the model simulations, sediment PCB concentrations change relative to the initial concentrations. These differences over time are caused by differences in the spatial patterns of erosion and deposition and are themselves caused by spatial variations in bathymetry and temporal variations in river flow.

Thus, given the fact that the very broad spectrum of PCB sediment concentrations has been present in the river for more than 30-40 years despite continuous sediment erosion and deposition over these years, it is clear that river erosion and deposition processes actually at work do not smooth the PCB distributions. Absent any explanation of how these PCB concentration distributions are developed and maintained over such a long period of time it seems very unlikely that the modeling will provide any believable answers to remediation strategies. This point has been made by me and others previously. Nevertheless, EPA and its contractors still say they do not know what processes are involved with the PCB fate and transport that would give rise to the very wide concentration spectrums observed. However, it is conceivable that the non-uniform erosion and sediment sorting processes not included in the model are indeed responsible for the non-uniform PCB distributions.

Response 2B-JL-4:

EPA would like to clarify the different spatial scales referenced by the Reviewer. Variations in PCB concentrations of two orders of magnitude were observed within areas of 1 m^2, with no variation in factors that influence erosion or deposition (e.g., depth, velocity, bed composition). Despite significant efforts, EPA was not able to identify the cause of this small-scale variability. EPA also recognized that it was not necessary to identify the mechanistic explanation for this small-scale variability to achieve the goal of the modeling study, where the scale of interest is larger than a grid cell and smaller than a reach. The decision was made to develop initial conditions taking this small-scale variability in the data into account. Spatial bins of approximately 400 meters in length (0.25 mile) were specified according to the procedures described in Appendix B.3 to the FMD. Assigning a single average concentration to all grid cells in a spatial bin, by definition, reduces the variability of the PCB concentrations represented in the initial conditions. Although cell-to-cell variations in PCB concentrations developed in the model over time due to cell-to-cell variations in erosion and deposition (MVR Figure 4.2-79), it is completely expected that the simulated concentrations have less variability than the original data used to calculate the average initial conditions.

This approach is consistent with the Reviewer’s comments from the MFD Peer Review (April 24, 2001):

The model that is proposed will attempt to describe the rates of erosion and deposition on a 20-meter grid plan. As the field and laboratory data collected by the Corps of Engineers would strongly suggest, predicting the rates of erosion on a grid this small cannot avoid the substantial error that is associated with the heterogeneity of the sediments. 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. …

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. However, 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…*
>
> Although the specific details of PCB release are not well-characterized (in terms of the timing, sizes, and mechanisms of release), the fundamental processes impacting sediment transport processes (erosion and deposition) are well understood. This understanding was central to model development efforts and permits successful simulation of how PCB exposure concentrations respond to changes in inputs over time. See FMD Sections 3 and 4.

I recognize the fact that these a very difficult problems to address, but if the modeling is to provide believable predictions of the future fate of the PCB in the sediments it seems to me that the processes that have resulted in the existing PCB concentration distributions have to be understood first.

> **Response 2B-JL-5:**
>
> As discussed previously, the processes that produce two orders of magnitude variations in PCB concentrations on a spatial scale of 1 m^2 are not resolved by the model, but because this spatial scale is far smaller than the scale of interest for the CMS, that level of spatial resolution is not needed to meet the goal of the modeling study.

2.3 GENERAL TOPIC 2C: SEDIMENTATION RATES

2.3.1 Comment Summary

Reviewer Name	Page Number/Line Number
Endicott	Page 6, line 17 to page 7, line 7
	Page 8, line 22 to page 9, line 3
Garcia	Page 3, line 35 to page 4, line 16

2.3.2 General Topic Summary

Net sedimentation is an indication of the difference between gross erosion and gross deposition over time. Because PCBs readily partition to solids, PCB concentrations in surface sediment are expected to respond to the net sedimentation over time. Consequently, the model response to changes in PCB inputs over time is strongly influenced by the ability of the model to accurately simulate net sedimentation rates.

One Reviewer noted that sedimentation rates in individual grid cells, particularly in Reach 5A, vary considerably and that there are no data that can be used to validate that result. Two Reviewers noted that the model-data comparisons of deposition in Woods Pond, where such data are available, suggest that the simulated sedimentation rates appear to agree well with the data.

2.3.3 General Topic Response

Although significant variations exist, EFDC is able to adequately simulate downstream trends in bed elevations over time on a subreach basis. Although cross-channel differences in erosion and deposition patterns are highly variable, these differences can be accommodated by aggregating the spatial and temporal scales of sediment transport processes. With process aggregation, a single model grid cell across the channel should be adequate to simulate sediment transport in the Housatonic River system. Other processes such as meander development and planform adjustment can be neglected as long as sediment inputs from bank erosion are explicitly included in the model.

The ability of the model to successfully simulate net sedimentation rates demonstrates that the mass of solids entering the model from all sources is correctly routed though the model domain. This is of particular importance for PCB transport simulations because the slow rates of net sedimentation in Woods Pond suggest that system response times to changes in PCB inputs will also be slow. The slow rate at which surface sediment PCB concentrations in Woods Pond change over time (the rate of PCB concentration change cannot be distinguished from zero) is further evidence that this important model construct is correct (see Appendix B.6 to the FMD and Section 4.3 of the FMD).

2.3.4 Responses to Specific Comments on General Topic 2C

Douglas Endicott

Sedimentation rates predicted by EFDC for individual grid cells in the river channel over the 10-year Phase 2 calibration period are plotted in Figure 4.2-58. This figure shows a highly variable pattern of deposition and erosion, especially in Reach 5A, where deposition fluxes exceeding 5 cm/yr are predicted for cells in close proximity to others in which erosion fluxes of 1-2 cm/yr are predicted. Predicted sedimentation rates are not tested by any data at this scale, however. On a reach basis, relatively small deposition fluxes (0.2-0.5 cm/yr) are predicted throughout the ROR for both the calibration and validation periods. Figure 4.2-60 graphs the sedimentation rate in grid cells within Woods Pond, along with sedimentation rates measured in cores. There are many EFDC grid cells in Woods Pond, and a significant number of dated sediment cores. I specifically examined the 1995 sediment core data suggested to be most representative in terms of cesium-137 profiles, and found that both measured and predicted sedimentation rates are in the range of 1 to 10 mm/y. On a pond-wide average basis, sedimentation rates predicted for the calibration period agree fairly well with the sediment core data (Table 4.2-5). Over the 26-year validation period, the predicted sedimentation rates appear to be somewhat low in comparison to data (Table 6.2-7).

Response 2C-DE-1:

EPA agrees with the Reviewer’s comments on the comparison of simulated versus measured sedimentation rates in Woods Pond, with the exception of the final statement that predicted sedimentation rates for the validation period (FMD Table 4.3-7) appear somewhat low in comparison to the data. In that table, the average sedimentation rate at the location of the core samples averaged 0.49 cm/yr. Sedimentation rates derived from the cesium data averaged 0.42 cm/yr using the peak in the 1995 cores, and 0.57 cm/yr using the first occurrence in the 1995 cores.

Predicted sedimentation rates were also compared to measurements of bathymetric change made between 2000 and 2006 at a series of transects in Reach 5A (Figure 6.2-33). The measurements of bed elevation change (positive and negative) were consistently larger and more variable than the corresponding model prediction although, since the predictions and data were both spatially averaged, I have trouble interpretting the significance of this comparison. As with other EFDC results for the sediment bed, grid cell predictions were aggregated to mile or sub-mile intervals. At finer spatial scales, there appear to be problems with the fidelity of the model. For example, the transect data from location XS153 show several feet of deposition while EFDC predicts significant erosion at this location, both during calibration (erosion=1”) and validation (4”) periods. The transect data themselves (presented at the Document Overview meeting: *09 - Erosion& Deposition - Garland.pdf*) show relative nonuniformity in sedimentation both longitudinally (between transects) and laterally (within transects). To me, the data show that lateral resolution is warranted for hydrodynamics (shear stresses) and sediment transport. Other subgrid features (lateral variations in bathymetry, bends, trees in river) are also significant physical features, which are not represented in the model.

Response 2C-DE-2:

To evaluate this aspect of model performance, EPA believes it is necessary to recognize (1) the spatial scale of model resolution, and (2) downstream and cross-channel differences in channel configuration over time. The minimum spatial resolution of the model framework is the size of an individual EFDC grid cell (20 to 70 m). In contrast, transects are composed of individual measurements along a line.

From transect survey to survey, measured sediment bed elevation changes for each subreach are variable. Differences between cross section average elevations indicate that net erosion occurred at some locations while net deposition occurred at others. When averaged over space and time, simulated changes in bed elevation are often near zero.

Cross-channel profiles are of interest because the patterns of erosion and deposition provide a means to assess the extent to which cross-channel differences in sediment transport processes exist and require aggregation. In many cases (21 of 53 profiles), sediment transport processes can be represented with a single grid cell across the channel without aggregation or significant loss of detail. However, in other cases (20 of 53 profiles), the profiles are so diverse that process aggregation would still be required even if three grid cells across the channel were used in EFDC. Given the need to integrate the floodplain and river channel into a single, computationally feasible model grid, little benefit is gained from adding model resolution to the river channel because aggregation of processes would still be required. See Appendices B.1 and B.6 of the FMD.

I remain concerned that we may be falsely confident in the erosion/deposition predictions made by this model, because the model is being applied and validated at relatively coarse spatial scales. Due to practical constraints, the modeling team did not simulate sediment transport at the resolution demonstrated in other successful sediment transport applications (Thompson Island Pool: HydroQual, 1995; Watts Bar Reservoir: Ziegler and Nisbet, 1995; Fox River: Lick et al., 1995; Saginaw River: Lick et al, 1995 and Cardenas and Lick, 1996; Buffalo River: Lick et al., 1995 and Gailani et al., 1996).

Response 2C-DE-3:

> The Reviewer acknowledges practical constraints that influenced the decision to adopt a grid scheme with generally one cell across the river channel. The factor that distinguishes the Housatonic River from the other systems is the need to include the floodplain in the model domain because approximately 60% of the PCB mass in the PSA resides in the floodplain. This presents a practical constraint to the modeling analysis.
>
> While features such as a meandering thalweg, sub-grid-scale undulations in bathymetry, and large woody debris affect the ability of the model to reproduce bathymetric changes at an individual transect, the performance of the model on broader spatial scales (e.g., sedimentation rates in Woods Pond, slow temporal changes in simulated PCB concentrations consistent with the results of a rigorous statistical analysis, and deposition patterns on the floodplain consistent with PCB concentration distributions) provides evidence of acceptable performance at the spatial scales that are relevant for the intended use of the model (see Appendices B.1 and B.6 of the FMD).
>
> However, it is important to recognize that the literature cited by the Reviewer is not directly applicable. None of the systems examined in these references address contaminant transport in the floodplain. A particular aspect of PCB transport and fate in the Housatonic River system is the ongoing, active (bi-directional) transport of PCBs and solids between the river channel and floodplain.

The unstated assumption is that applying a similar model on a cruder resolution will correctly average the sub-grid scale phenomena, even though some of these phenomena are demonstrably nonlinear. This assumption is not supported by either data comparisons or numerical convergence testing using alternative grid resolutions. Predictions of net sediment accumulation can only be tested for Woods Pond, because that is the only portion of the ROR where a reach-wide average accumulation could be inferred from sediment data. Predictions of net sediment accumulation are untested in other reaches. This is of more than academic interest, because we are asked to believe in the model primarily on the basis of the water column predictions of TSS and PCB concentrations. Yet these only indirectly measure erosion and deposition. Reach 5A, for example, is described as a dynamic environment with rapid exchange of solids and associated PCBs. Within this and other reaches, there are locations where both deposition and erosion are taking place. If aggregated spatially (into either reaches or miles) much of this behavior is “averaged out”. Aren’t there advantages to validating sediment transport predictions without this spatial averaging? In fact, that is the only way it can be done.

Response 2C-DE-4

EPA has demonstrated that the model framework correctly simulates water column TSS and PCB concentrations during storm events, long-term net sedimentation rates, and surface sediment PCB concentration trends over time in Woods Pond. This confirms that the balance of fate processes in the model is correctly represented.

Model-data comparisons of TSS concentrations during storm events provide more value than is recognized by the Reviewer. Because TSS data were collected at a fine temporal frequency, the data provide a description of conditions on the rising limb of the hydrograph, when little deposition is expected, and on the falling limb, when deposition dominates the solids transport processes. A sensitivity analysis presented in the Model Calibration Responsiveness Summary (page 23) demonstrates that the balance between erosion and deposition is constrained by the gradient between the water column and bed PCB concentrations. The averaging of sub-grid scale variations in bathymetry, shear stress, and PCB concentrations, therefore, has been demonstrated to produce an adequate representation of the average response on water column TSS and PCB concentrations.

The Reviewer correctly notes that Woods Pond is the only location where net sedimentation can be adequately tested through comparisons with cesium-137 data. Simulated net sedimentation rates for Woods Pond are in close agreement with measured values. This close correspondence is significant because it demonstrates that on average the mass of solids entering the model from all sources is correctly routed though the model domain (including deposition on the floodplain and in the river channel upstream of Woods Pond). This means that sediment input from each source (upstream, tributary, sediment bed, and river banks) is correctly transported through the river and floodplain such that the cumulative mass of solids deposited in Woods Pond over time matches the independent estimates of net sedimentation. The slow rates of net sedimentation in Woods Pond suggest that system response times to changes in PCB inputs will also be slow. This is demonstrated by the fact that the rate at which surface sediment PCB concentrations in Woods Pond change over time is slow (i.e., the rate of PCB concentration change cannot be distinguished from zero) (see FMD Appendix A.1). In essence, this is a straightforward application of a basic engineering principle, conservation of mass (continuity).

Because the intended use of the model is at a resolution larger than a grid cell but smaller than a reach, validation of the model at the grid scale and subreach level is the appropriate scale for demonstrating model performance. Therefore, EPA disagrees with the Reviewer’s premise that the only (implying the appropriate) scale for model validation is at a subgrid scale.

Marcelo H. Garcia

Of particular relevance for this reviewer is the validation of flood plain deposition throughout the river system as well as sediment erosion and deposition in Woods Pond, since the record shows that this pond is a major deposit for PCB-contaminated sediments.

Regarding the sedimentation of Woods Pond, the calibration/validation exercise has produced results showing rates of (*net*) sedimentation for the period going from 1979 to 2004 (Figure 6.2-34). The results seem reasonable with predicted sedimentation rates (mm/year) that are within an order of magnitude of sedimentation rates determined from Cesium data. The model also reproduces fairly well the distribution of PCBs with depth as observed from sediment cores. However, as stated below, the model shows some bias when predicting temporal variations of surface PCB concentrations in Woods Pond.

Response 2C-MG-1:

EPA agrees with the Reviewer's comment that the comparisons between measured and simulated sedimentation rates are reasonable, and points out that the agreement on a core-by-core comparison is better than the order-of-magnitude characterization mentioned by the Reviewer. A more detailed statistical analysis of temporal distributions of PCBs in Woods Pond sediment has been completed and is presented in Appendix A.1 to the FMD. The results of this analysis indicate that the Reviewer's interpretation of a bias in the prediction of temporal variations in surface PCB concentrations in Woods Pond is incorrect.

Woods Pond is perhaps one of the few locations in the Housatonic River where the capabilities of the model to predict the spatial and temporal distribution of sediment and PCB can be extensively tested due to the large amount of observations available. This has been done to some extent but more effort should go into this since the pond could become a major source of PCBs to the downstream portion of the river during a major flood event.

Response 2C-MG-2:

Additional comparisons between simulated and measured sedimentation rates in Woods Pond, which were not included in the MVR, have been included in the FMD (Section 4.3 and Appendix B.6). A more detailed statistical analysis of temporal distributions of PCBs in Woods Pond sediment has been completed, and is presented in Appendix A.1 to the FMD.

3. GENERAL TOPIC 3: VERTICAL AND HORIZONTAL SPATIAL RESOLUTION

3.1 GENERAL TOPIC 3A: NUMBER OF GRID CELLS ACROSS CHANNEL

3.1.1 Comment Summary

Reviewer Name	Page Number/Line Number
Adams	Page 12, lines 18 to 35
	Page 13, lines 1 to 18
Bohlen	Page 8, lines 1 to 15
	Page 10, lines 35 to 36
	Page 14, line 32 to page 15, line 15
Endicott	Page 2, lines 18 to 19
	Page 2, line 25 to page 3, line 5
Garcia	Page 2, lines 20 to 31
	Page 2, line 37 to page 3, line 6
	Page 5, lines 4 to 9
	Page 5, line 39 to page 6, line 2
Gobas	Page 4, lines 17 to 31
	Page 4, lines 38 to 45
Lick	Page 1, lines 19 to 28
	Page 11, line 20 to page 12, line 17
	Page 13, lines 16 to 17
List	Page 2, lines 4 to 42
	Page 3, lines 22 to 35

3.1.2 General Topic Summary

A grid must be specified for computer models of the environment to generate numerical solutions. A grid can be described as a collection of the points in space where model computations are performed in sequence over time. The process by which the physical region (domain) represented by a model is divided into points used for calculations is called grid discretization. Within each individual grid cell, the physical properties of the environment are represented as averages. There is a tradeoff between the number of cells in the grid (grid resolution) and the number of computations that must be performed (computational effort).

Six Reviewers questioned the representation of channel width by a single cell in the modeling grid, with most suggesting that three or more cells should be used across the channel. Most Reviewers noted that flow and velocity are variable across the channel and that erosion and deposition processes are also variable. In particular, Reviewers noted that sediment transport processes vary non-linearly with flow such that averaging across a single cell in the river channel does not provide a good simulation of processes known to be important for sediment and contaminant transport. Most Reviewers also recognized that additional grid cells would make the model more computer-intensive which is a concern.

3.1.3 General Topic Response

Selection of an effective model grid was particularly important for the development of the EFDC application to the Housatonic River because EFDC simulates PCB transport and fate at the smallest spatial and temporal scales that occur in the model framework. As described in FMD Appendix B.1, efforts to select a grid for EFDC focused on finding a balance between grid resolution and computational effort based on six factors, four that influence grid resolution at a conceptual level and two that reflect practical constraints on grid resolution that are imposed by the existence (presence or absence) of data and computational resources. These factors establish the balance between what might be possible and what can be accomplished in the context of the modeling study goal.

The model grid was chosen as a reflection of the nature of the Housatonic River system (channel and floodplain) and the modeling study goal. Much of the PCB mass in the system is in channel banks and floodplain areas. River flows periodically overtop channel banks and inundate the floodplain. This connection between the river channel and floodplain is important because risks to both human and ecological receptors were identified from exposure to PCBs in the floodplain. Further, the goal was to develop a modeling framework that can be used to adequately simulate baseline conditions and the relative performance of potential remedial alternatives. The spatial scale needed to adequately achieve this goal was defined by the Model Working Group as smaller than a river reach but larger than a single grid cell (QEA, 2001; EPA, 2006). The temporal scale of concern is on the order of decades, sufficiently long to evaluate the effectiveness of remedial alternatives against baseline conditions. Collectively, these considerations require that floodplain areas be explicitly represented in the model framework and that model results be accurate when evaluated on a subreach scale.

A number of different model grid configurations were evaluated before the final grid for EFDC was selected. Sensitivity analyses were conducted using detailed channel cross sections, bathymetry and velocity information collected for a 1-km subsection of the river during the Test Reach Investigation. Both in-bank and out-of-bank flow conditions were examined. An overview of the Test Reach Investigation was presented in the Section 5.4 of the Model Framework Design. A further summary is presented in FMD Appendix B.1.

The need for explicit floodplain simulation imposes an important constraint on the type of model grid that can be used. In an ideal simulation, the grid would allow high resolution in the river channel to simulate differences in conditions across the channel because such differences can influence the extent of bank and bed erosion. However, even if this resolution was added to a model grid in the river channel, the model results would not necessarily be more accurate. Moreover, this degree of detail in the channel would be needed only if the model framework were going to be used to simulate sediment and PCB transport at the scale of individual bars or individual grid cells. In the application of the model framework for the Housatonic River, much larger spatial scales are of concern, thereby reducing the need to simulate cross-channel variations. This means that the effect of fine-scale processes can be adequately represented in a simplified, aggregated fashion. Further, at the scale of fish movement (FCM reaches) and risk assessment output, model results are not expected to be sensitive to use of a more finely resolved model grid.

The evaluation of potential remedial alternatives is expected to focus on relatively large spatial units (larger than a grid cell) and that fine-scale details, even at the scale of individual grid cells, would always be aggregated and averaged regardless of the resolution of the underlying model. In addition to information EPA provided to the Reviewers, the need for process aggregation using a less detailed model grid was presented by GE during their public comment to the Model Framework Design Peer Review (see QEA, 2001). Moreover, even if the spatial scales of interest were ignored, it is worth noting that cross-channel conditions are sometimes so variable that process aggregation and averaging would still be needed even if the river were represented using three grid cells across the channel. This type of cross-channel variability occurred in 20 of 53 profiles examined in FMD Appendix B.6. Consequently, given the intended use of the model framework and the potential for cross-channel variation even at extremely fine spatial scales, some degree of process aggregation and averaging will be needed regardless of how the river channel is represented in the model.

In consideration of all the factors that influenced selection of the model grid for EFDC, a curvilinear grid that includes the 10-year river floodplain and represents the channel width using a single grid cell was judged to provide the best balance between spatial detail and computational efficiency as needed to meet the goals of the Housatonic River modeling study.

3.1.4 Response to Specific Comments on General Topic 3A

Eric Adams

Grid Resolution

Like most other review panelists, I continue to be **concerned about the lack of lateral resolution in the computational grid**. Having only about one grid cell over the river width means that important processes must be parameterized. Because the existence and magnitude of erosion and deposition depend non-linearly on stream velocity, a trapezoidal section with depth varying linearly between 1 and 3 meters at the two banks will look to the model like a section with uniform depth of 2 m, yet in practice the former may have regions of strong erosion and deposition (as in re-surveyed cross-section XS061 shown in the DRM handouts), whereas the latter might be marginal one way or the other.

Response 3A-EA-1:

Given the intended scale of use for the model framework, EPA believes that the net transport of solids and PCBs through the water column and average PCB exposure concentrations are the important features of the river system that must be accurately simulated on a subreach basis. As shown in FMD Section 3, the present model framework achieves this needed level of accuracy.

As noted in the General Response, some fine-scale detail is lost due to process aggregation. Specifically, cross-channel differences in erosion and deposition are aggregated such that the model response may seem less dynamic than the response that might be measured at any one point along or across the river. However, it should be recognized that the reduced precision in model results that arises from averaging does not mean that the results are less accurate.

First, process aggregation would be required in any model, regardless of scale, as a function of the intended scale of model use and point-to-point variability that occurs in the Housatonic River. Second, for any subreach, the flux of solids and PCBs moving through the water column is known from measurements. This means that the net flux solids and PCBs transported from any subreach can be determined with reasonable accuracy. It must also be recognized that the sediment bed of the Housatonic River is heavily contaminated with PCBs across the entire bed and that pockets of uncontaminated sediment do not exist. Consequently, it is not necessary to know the exact part of a model grid cell (or even exactly which cell) from which sediments and PCBs are eroded. Even if the model does not identify the precise location from which erosion occurs, the model still accurately simulates the average flux through the system because those fluxes are determinable from site-specific measurements and are accurately matched. Thus, the model is accurate even if some precision is lost through averaging.

As noted in the General Response above and in FMD Appendix B.6, there are many locations (20 of 53 profiles examined) where cross-channel differences in erosion and deposition are so extensive that processes would need to be averaged and aggregated even if the river were simulated as three cells across the channel. However, as shown in FMD Appendix B.1, the computational burden of using multiple cells across the channel is prohibitive. Given the computational burden and the need to aggregate processes at some level regardless of model resolution and inclusion of the floodplain, EPA believes that representation of channel processes with a single cell across the river is appropriate.

Process Aggregation Prevents Evaluation of Remediation

Even if the model gets the net erosion/deposition correct (doubtful due to the non-linearity), the failure to capture gross erosion/deposition is problematic, both for evaluating remedial measures (such as capping) and natural attenuation. Assume an entire channel cross-section is capped. If half of the channel is erosional and the other half is depositional, the first half will require added protection to keep the cap intact, while the second half will get covered with contaminated sediments from upstream. Neither process would be forecast with a model that predicts marginal net erosion/deposition. As for natural attenuation, contaminated sediments from an upstream area with gross erosion can contaminate (or re-contaminate) downstream areas, but the model would not predict this.

Response 3A-EA-2:

As noted in the previous response and as discussed in Sections 3 and 4 of the FMD, the present model framework achieves the level of accuracy needed for the modeling study. Further, it should be noted that the function of the model framework in the next phase of the project will be to provide an estimate of the relative difference between remedial alternatives.

The Reviewer's comment regarding the use of the model framework for evaluating remedial alternatives does not accurately reflect the intended use of the model. As noted, the model will be used to assess the relative difference between simulations. Specifically, the impact of any given remedial option can be successfully represented by altering model boundary conditions and initial conditions to simulate the impact that the management of contaminants in Housatonic River system will have on decadal time scales. The model is not intended to be used a tool to engineer specific components of the selected remedial alternative. Therefore, if capping were selected as a proposed remedy for a portion of the site, guidance and procedures specific to cap design would be used (e.g., USACE guidance for cap stability).

Therefore, the model framework must be considered a tool to evaluate the relative impact that a given remedial strategy (such as a capping) is expected to have on a level larger than a grid cell and smaller than a reach (e.g., the FCM reaches). The model can serve this intended purpose by use of altered boundary conditions and initial conditions to construct difference scenarios representing different management strategies. The model framework will not be used to address the engineering issues associated with the design and implementation of a remedy.

Scale-dependence of EFDC Calibration

Too begin with, it must be recognized that the lateral grid size is a model parameter as well as a numerical parameter. Ideally, when a set of equations are solved using a discretized numerical model, one likes to reduce the grid size until results converge. ***But, we are no where near that here.*** Hence if the model grid size were to be reduced, this would affect other model processes (notably erosion, which depends heavily on velocity, which would vary significantly across the river width). Thus ***if the model grid were to be reduced, the entire calibration would need to be redone.***

It is not clear how much resolution would be needed to properly compute transport. Because of this uncertainty, and the time and effort involved, it may be too late in the game to make major changes to the model grid. However, at a minimum, some sensitivity tests could be conducted using a finer grid over a small portion of the (in-channel) domain and run over a short duration of the 26 year simulation period. The output from the more highly resolved model could be compared with that from the coarser model and then parameterized. For example, it might be determined that the net erosional flux with a resolved grid is X times that computed with the coarse grid, in which case predicted erosional fluxes with the coarse grid would be multiplied by X.

Response 3A-EA-3:

EPA recognizes the need for sensitivity analyses and the potential grid-scale dependencies that may exist in the model framework calibration. For EFDC, EPA presented a series of analyses in which model sensitivity to downstream and cross-channel grid resolution in the channel was examined. These sensitivity analyses were first presented in the Final MFD and are again summarized in FMD Appendix B.1. Stream velocity as calculated by the hydrodynamic model was the main parameter examined. EFDC results for Cartesian and curvilinear grids of different resolutions were compared. See Tables 1 and 2 of FMD Appendix B.1.

Further, it should be recognized that some degree of grid scale-dependency will exist in any practical model application. As model scale varies, processes that might have been negligible at one scale may become predominant at a different scale. This is why it is important to understand the intended spatial and temporal scales for which the model will be applied. Moreover, the difference between what happens at a single point can differ from what is observed over a broader area. This was well-documented throughout the modeling study as conditions in the Housatonic River system vary widely over very small spatial scales (just a few feet). Selection of any model spatial (or temporal) resolution requires that processes be aggregated and averaged. This is a fundamental concept in the difference between the integral and differential approaches in fluid mechanics.

The correct functioning of the model framework for the present grid scale, which was specifically selected to meet the goal of the modeling study, is well documented in Sections 3 and 4 of the FMD.

W. Frank Bohlen

Model-Data Comparisons and Grid Scale

Comparisons between measured and modeled velocities are complicated by the model use of a single grid cell across the main stem channel. If the single cell is to be retained measurements should be designed to yield cross-sectional averages as opposed to point measurements. The majority of available data do not appear to be suitable for this purpose.

Response 3A-FB-1:

Velocity measurements used for model-data comparisons are cross-section averages. Water quality samples are either depth-integrated (manual samples) or at a fixed depth (ISCO samples). As noted by the Reviewer, it might have been possible to facilitate model-data comparisons if all measurements were always cross-sectional averages.

Need for More Grid Cells Across the River Channel

While on the subject of model grid specification, I must repeat what has been stated before regarding the advisability of increasing the number of grid cells across the channel. My review of available data detailing bathymetry and sediment type indicates that accurate simulation of these characteristics and their proper incorporation within transport models requires a minimum of three grid cells rather than the present one. If this proves (or is known) to result in an unacceptable increase in computation time then consideration should be given a reduction in the lateral extent of the floodplain since there are a variety of data (see Example Model Simulations, e.g.) that suggest that there are substantial areas along the inshore limits of the floodplain that are only occasionally affected by flooding and where, as a result, minimal longterm changes are to be expected. Alternatively, a coarser grid might be used on the floodplain.

Response 3A-FB-2:

Use of a coarser model grid for the floodplain does not alter the computational burden because: (1) model time steps would still be controlled by downstream and cross-channel velocities in the river channel; and (2) model computations do not occur in floodplain cells when those cells are not flooded. See General Response above and Appendices B.1 and B.6 to the FMD.

Examination of these data might also be part the studies dealing with the recommended increase in the number of model grid cells across the river channel.

Response 3A-FB-3:

The choice of a single grid cell across the river channel was driven by the balance between the intended use of the model and the corresponding computational burden of added resolution. While data collection at a finer scale might prove to be necessary for design of a remedy, EPA believes that no further data collection is needed to meet the goal of the modeling study.

Even with these process questions resolved there remains the issue of run-time. It seems clear that the model as presently configured requires entirely too much time for the completion of a single run to be useful within timely evaluations of a significant number of remedial schemes. We probably knew this several years ago and should have been more sensitive to the need to develop alternative formulations. A number of these, including the separation of the hydrodynamics from the transport estimates and subsequent FCM evaluations were previously mentioned. It is now necessary and possible to go further. Using the experience gained from “whole model” runs it should now be possible to develop a number of synthetic hydrographs detailing streamflow, stage, and TSS concentrations at the upstream boundary and each of the primary tributary streams. This would eliminate the need to run HSPF on a regular basis. EFDC is a ponderous model and can be streamlined now that we have a better idea of the relative importance of the governing variables. John Hamrick should be charged with this task (no more than 6 months) as soon as possible. As part of this streamlining the model grid characteristics should be carefully reviewed, again using what has been learned about the relative importance of each of the domain regions (sidebanks, backwaters, floodplains etc.). My sense of the present grid is that it underspecifies the channel region and overspecifies the floodplain. The latter could almost be treated as a box with fluxes simply proportional to stage which could be specified along its margin by EFDC. If in time more detail of the interior of the plain is needed for remedial purposes consideration might be given to replacing the high resolution grid on the plain while placing a box in some other area (channel sidebanks?). Finally, I'd consider eliminating the FCM in favor of a parametric (flow, TSS concentrations?) approximation of body burden uptake based on the results of the complete model runs.

Response 3A-FB-4:

Model framework run times are controlled by the time required to run EFDC. No meaningful reduction in run times can be accomplished by altering HSPF or FCM or the linkage between these tools and EFDC (see General Topic 4). As previously noted, coarsening of the EFDC grid for the floodplain will not significantly reduce run times because the model does not perform calculations on cells when they are dry and because model time steps for numerical integration are restricted by stream velocities in the channel.

Further, EPA has examined many approaches to reduce run times; over the course of the modeling study, the time required to complete EFDC simulations has decreased significantly. As of October 2006, the run time for a 1-year simulation is approximately 8 hours using a computer with two Intel Xeon 5160 Dual-core CPUs running at 3.0 GHz with 4 GB of RAM running the Mandriva Linux 2007 operating system for X86-64. Even more recent computers may have the power to further reduce these run times. See responses in General Topic 4.

Douglas Endicott

EFDC fails to address lateral variation in hydrodynamics and sediment transport processes

Response 3A-DE-1:

Lateral variations in hydrodynamics do not need to be simulated at the intended scales of use for the model framework. See the General Response above.

On the first point, EFDC appears to fail to address lateral variation in hydrodynamics and sediment transport processes, because the model grid in the river channel is one-dimensional over most of the Rest-of-River (ROR). Prior applications of SEDZL (the model upon which the EFDC sediment transport code was apparently based) to river systems have used a 2-

dimensional, vertically-integrated grid, using 5 to 10 lateral grid cells to represent the channel cross section. The 2-dimensional sediment transport models have been demonstrated to predict scour and deposition in river channels with reasonable accuracy[1]. These applications have demonstrated that lateral variation in current velocities and bottom shear stresses result in considerable variation in sediment scour depending upon location in the channel. This behavior is often observed in natural channels as a pattern of erosion from mid-channel/deposition nearshore, or erosion on the outside of a bend/deposition on the inside. Because the relationships between current velocity, shear stress and sediment erosion are nonlinear, it appears likely that sediment scour predicted by the 1-dimensional model of sediment transport in the Housatonic River will be different than the scour predicted by a 2-dimensional model. How much different is unknown, because the EPA modeling team have not tested the sensitivity of the EFDC sediment transport model to the number of lateral grid cells in the river channel. This is a significant issue, because release of PCBs from eroding sediments is a major source of the contaminant in the ROR.

Response 3A-DE-2:

See General Response above. In addition, at the intended scales of use for the model framework, cross-channel variation in sediment transport would be aggregated and averaged regardless of the resolution of the underlying model. It must be recognized that the sediment bed of the Housatonic River is heavily contaminated with PCBs across the entire bed and that pockets of uncontaminated sediment do not exist. Consequently, it is not necessary to know the exact part of a model grid cell (or even exactly which cell) from which sediments and PCBs are eroded.

The literature citing SEDZL applications to other sites is not appropriate to the Housatonic River system because none of those applications to other sites (such as the Lower Fox, Saginaw, or Buffalo Rivers, etc.) couple the river channel and floodplain into a single model domain or include any part of the floodplain. A distinguishing feature of the Housatonic River system is that 60% of the PCB mass inventory in Reaches 5 and 6 is present in the river banks and floodplain areas and that risks to human health and the environment are posed by exposure to PCBs in the floodplain. The overriding need to couple river channel and floodplain processes is a major constraint on the type of model framework that can be used for the Housatonic River system.

EPA presented a series of analyses for EFDC in which model sensitivity to downstream and cross-channel grid resolution in the channel was examined. These sensitivity analyses were first presented in the final MFD and are again summarized in FMD Appendix B.1. Stream velocity was the main parameter

---

[1] See: Gailani, J.Z., W. Lick, K. Ziegler, and D. Endicott. 1996. Development and Calibration of a Fine-Grained Sediment Transport Model for the Buffalo River. *J. Great Lakes Res.*, 22(3):765-778.

examined and EFDC results for Cartesian and curvilinear grids of different resolutions were compared. See Tables 1 and 2 of FMD Appendix B.1.

Marcelo H. Garcia

Arguably the main shortcoming of the model is not the Housatonic River model itself but rather its numerical implementation which, in my opinion, has rendered the model calibration and validation unreasonably difficult. Assuming that all the processes are eventually accounted for in the model and the right algorithms are used to simulate them, the most vexing issue will continue to be the characteristics of the computational grid used to simulate the river hydrodynamics, sediment transport and associated PCB transport and fate. The computational mesh (Figure 3-6) is "hard-wired" to water stages in the floodplain associated with a flood having a recurrence time of 10 years. This results in the main river channel being modeled with the same spatial resolution (i.e. a single computational cell) regardless of the flow discharge, not just the 10 year flood.

Response 3A-MG-1:

See General Response above. It should again be noted that the model grid resolution across the channel is appropriate because finer detail would be aggregated and averaged at the intended spatial and temporal scales of use for the model framework.

At the same time, the use of a course grid hinders the ability of the model to resolve the flow velocity distribution inside the main channel as well as the associated boundary shear stresses throughout its wetted perimeter. Given that erosion and sediment transport vary non-linearly with increments in flow velocity and shear stress, it is rather imperative to find a way to increase the spatial resolution of the model so that at least three "streamtubes" are used to model flow and sediment transport in the main channel of the Housatonic River. This will in turn facilitate and make more meaningful the computation of streambank erosion as well bed sediment erosion and resuspension.

Response 3A-MG-2:

It should be recognized that streambank erosion is parameterized as a load using site-specific empirical functions that accurately reproduce long-term and short-term erosion rates measurements for the Housatonic River as a function of cross-section average flow stream (see Appendix B.7 to the FMD). Use of these empirical functions eliminates the need to simulate cross-channel variations in velocity and shear stress along the channel banks with extreme detail. Furthermore, as demonstrated in Sections 3 and 4 of the FMD, the model is able to correctly resolve average sediment transport processes over the spatial scales of interest.

As mentioned in the response to question one, the size of the computational grid makes it difficult (and meaningless) to compare model predictions for processes that take place at hydraulic and sedimentation scales determined by the flow rate and the main channel width.

This is not the case for Woods Pond and the floodplains were the size of the computational grid is appropriate for the scale of the flow field.

Response 3A-MG-3:

Woods Pond introduces a different flow regime than exists in Reach 5. Two-dimensional representation of the river flow in Wood Pond is appropriate because cross-pond differences in flow in this reach are significant for simulating sediment and PCB transport at the spatial scale of interest to the modeling study. However, in Reach 5 and free-flowing areas in Reaches 7 and 8, simulation of the channel as a cross-section average is appropriate and justified given the intended use of the model framework and the fact that downstream flow velocities will always be much larger than cross-channel velocities in these areas.

Given the facts that several processes are yet to be fully characterized in the model and that a coarse computational grid has been used for the model calibration, I do not think that a full sensitivity analysis can be conducted at this time.

Response 3A-MG-4:

Model sensitivities to grid resolution have been characterized. See Response 3A-EA-3 above.

Frank Gobas

The spatial resolution of the environmental fate and food-web model is unnecessarily complex. A much simpler spatial model is consistent with the empirical data and the current model calculations. The hydrodynamic model does not run on a low spatial resolution. However, it is possible to run the hydrodynamic model on its optimal spatial resolution and aggregate the hydrodynamic data for a much simpler PCB environmental fate model. The latter would also reduce the model run-time from the current, unacceptable 30 to 50 d. This would also make the model more transparent due to greater simplicity. The modelers could argue that in terms of predicting spatially varying concentrations the current multi-compartment model is consistent with a much simpler lower spatial resolution model, so there is no need to develop the simpler model. This is true for the model’s current application. However, the modelers should keep in mind that when the current model is applied to make predictions of the impact of remedial actions on PCB concentrations as a function of space (i.e. model objective #1), the model may predict spatial differences in concentrations that have not been “validated” or “ground truthed”.

Response 3A-FG-1:

EPA believes that the spatial and temporal resolution of the model framework, particularly the grid for EFDC, is appropriate for the scales of interest for the modeling study. However, as noted in Response 3A-FB-3, model run times are not expected to be overwhelming because the run time for a 1-year simulation is approximately 8 hours. This means 50-year long simulations can be completed in 17 days.

The Model Working Group will continue discussions regarding various levels of coupling/decoupling of model results as part of the Corrective Measures Study. However, such decoupling was not needed for the development of the model framework as conducted by EPA. See also the responses in General Topic 4.

The modelers should consider developing a simpler spatial representation of the environmental fate and food-chain model by decoupling the hydrodynamic and sediment transport model from the environmental fate and food-chain model. The current high spatial resolution of the model is unnecessary for the environmental fate and food-chain model. Also, the time-step requirement for the hydrodynamic model is too onerous and unnecessary for the environmental fate and food-chain models. This change in model design will make it possible to make many model runs within a reasonable computational time.

Response 3A-FG-2:

See Response 3A-FG-1 above.

Wilbert Lick

In the present model, the processes of sediment erosion, sediment deposition, the finite sorption rate of PCBs by the sediments, and the sediment-water flux due to “diffusion” are described incorrectly and inaccurately. This is exacerbated by a very coarse numerical grid used to define the bathymetry of the river and an unnecessarily fine grid to define the bathymetry/topography of the floodplain. More specific reasons for these comments as well as specific suggestions for improvement of the model are presented in the complete response which is attached. I do not believe that the present model can reasonably account for the relevant processes affecting PCB fate, transport, and bioaccumulation in the Housatonic River to a degree consistent with achieving the goal of the modeling study.

Response 3A-WL-1:

EPA believes the Reviewer comment is incorrect, particularly in regard to the balance of fate processes in the model framework (see General Response in General Topic 1). With respect to the model grid, explicit coupling between the river channel and floodplain is required because roughly 60% of the PCB mass inventory in Reaches 5 and 6 is in the floodplain (including channel banks). The need to include floodplain areas constrains the model grid that can be used. Furthermore, there will need to be some process aggregation and averaging of model results regardless of whether the river is represented as 11, 5, 3 or just 1 cell(s) across the channel.

As far as the high PCB variance throughout the river is concerned, consider the following. Sediment erosion/deposition depends on the hydrodynamics, e.g., high rates of erosion where the flows are fastest and low rates of erosion (or deposition) where the flows are slow. Because of this, large variations in erosion rates occur across the river (shallow, near-shore areas versus deeper channels in the middle) as well as along the river. This is well illustrated in previous calculations of sediment transport in rivers (Saginaw River (Cardenas et al. 1995), Fox River (Jones et al. 2000)) when a reasonably fine grid was used, i.e., 5 to 11 grid points across the river. Because of the coarse numerical grid, this is not described by the present model. i.e., everything in the model is averaged or smoothed. In some areas (e.g., upstream and/or in the center of the channel where flows are high), mostly erosion occurs and PCB concentrations reflect deposition at an earlier time and are relatively high. These sediments will be transported downstream and will deposit in a non-uniform manner depending on the hydrodynamics. These sediments will retain their high PCB concentrations. In other areas (e.g., sediments from above the upstream boundary, near-shore depositional areas which have received clean sediments from further upstream, or areas where deposition is slow such that surficial sediments can equilibrate with the cleaner overlying water), the PCB concentrations of surficial sediments may be quite low. Slow transport and multiple resuspension/deposition events can also cause low PCB concentrations of surficial sediments.

Because of episodic events, the dependence of erosion/deposition of sediments on highly variable hydrodynamics, the highly variable sources of PCBs (e.g., clean from far upstream, contaminated from deep in the sediments, differences between near-shore, shallow areas and the deeper channel in the center), and of course slow PCB desorption rates (which causes sediments to retain their PCB sorbed concentrations, either high or low), it seems quite plausible that PCB concentrations will be highly variable in space and time in the Housatonic. The present model smooths this all out.

Does this matter? If the only purpose of the model is to duplicate known results, then accurate models of sediment and contaminant transport and fate don't matter. However, if the model is to be used for predictive purposes, then accurate process models do matter. In the predictive mode, future conditions (such as sediment properties, contaminant concentrations in the sediments, concentrations and types of benthic organisms, sediment-water fluxes, flow rates, etc.) will be modified (for example by dredging, capping, or extreme environmental conditions), will change with time, and will be different from those for which the model was calibrated. The basic processes in the present and future are the same. However, their relative effects and significances depend on the modified conditions and will change with time. If the models describing the basic processes have incorrect functional behavior and/or inaccurate parameters, then the model will not predict the long-term behavior properly. Because of this, for the long-term prediction of sediment and contaminant fluxes, it is essential that the functional behavior and parameters of the most significant processes in the model be described correctly.

Response 3A-WL-2:

The Reviewer comments intertwine the issue of model grid resolution with an assessment of the balance of fate processes in the model framework (see General Topic 1 for responses regarding the balance of processes). With respect to the resolution of the grid, the Reviewer comment places a greater emphasis on cross-channel differences rather than downstream gradients. To achieve the goal of the modeling study, the model framework must be able to describe PCB transport and fate over relatively large spatial units over decadal time scales. As documented in Sections 3 and 4 of the FMD, the model framework successfully defines the relationships between PCB sources and resulting concentrations with appropriate accuracy.

The numerical gridding can be improved by increasing the number of grid cells across the river and decreasing the number of grid cells in the floosplain.

Response 3A-WL-3:

See the General Response above, and Appendices B.1 and B.6 to the FMD.

E. John List

EPA and its contractors decided to finesse the problems associated with the non-uniformity of stream bed sediment erosion by using a single computational cell to define the entire width of the river. This means that the bed erosion at any specific river station is defined by the mean bottom shear stress. However, because the actual shear stress on the bed at any location is proportional to the local *depth-averaged* velocity, and the rate of bed erosion (when the shear stress significantly exceeds the critical shear stress) is proportional to the shear stress squared (or an even higher power), it means that the rate of bed erosion in these circumstances is proportional to the fourth power (or even higher) of the mean velocity. Thus, if there is a transverse non-uniform depth-averaged velocity profile *across* the river the rate of bed erosion may vary quite substantially with location on the river cross section. At locations where the depth-averaged velocity is twice that at another location the rate of erosion can be 16 times as great (or even higher). The fact of the matter is that real streams do have widely varying flow velocities across the width of the river and this is the basic reason why the water depth is seldom uniform across the width of the river. It is my professional opinion that it is really inappropriate to attempt to describe the net result of the widely varying rates of erosion that occur in the Housatonic River by using a *cross-sectionally* averaged velocity. The probability that the mean erosion can be parameterized in the model on the basis of cross-sectionally averaged flow properties seems low to me. However, development of a sediment rating curve from the stream data may show that this is possible (see discussion below).

It is acknowledged that it is probably quite impracticable to model the sediment transport and at the same time compute the resulting changes in river bottom topography and their effects on the hydrodynamics. It would seem that a logical approach would be to use a quasi-steady approach in which the stream bed is divided transversally into three or more computational elements that enable at least a partial simulation of the bottom profile and the resulting velocity distribution and bed erosion. An alternative would be to take a longitudinal section of the stream and do a detailed three-dimensional computational fluid dynamics (CFD) analysis of the flow distribution and resulting erosion, then compare the results of this analysis to the one-dimensional approach to determine if a realistic one-dimensional parameterization is at all possible. Perhaps the single best alternative approach to this problem is simply to use empirical sediment and flow rating curves, which is the "old fashioned" empirical way these problems were solved before computers became involved (see, for example, http://www.epa.gov/warsss/sedsource/rivrelat.htm and http://water.usgs.gov/osw/techniques/TSS/Horowitz.pdf, where it is shown that accuracies of

±20% are possible for empirically-derived sediment concentration predictions). Incidentally, the hydraulic modeling could probably also benefit from a comparison of the predicted and actual stage-discharge curves, if this has not already been done.

Response 3A-JL-1:

As noted in the General Response above, some degree of process aggregation and averaging is necessary regardless of whether the river is represented as many cells across the channel or just one. Nonetheless, the model framework can simulate average sediment transport conditions, as demonstrated in Sections 3 and 4 of the FMD.

However, it should be noted that the alternative approaches suggested by the Reviewer would not be expected to materially improve model performance. The use of empirical rating curves is in many regards no different from the present model calibration approach. Further, the use of quasi-steady flow simulations to characterize hydrodynamics may be a less reasonable approach because river flow is known to be unsteady (vary over time). Temporal variations in flow cause variations in flow accelerations and corresponding sediment and PCB transport. EPA believes that accurate hydrodynamic simulation requires that the model accurately simulate these transient dynamics.

A second reason for this opinion is that recent river bottom survey profiles indicate quite clearly that sediment erosion and deposition is very definitely non-uniform across the river cross-section. By contrast, and as discussed above, the modeling uses a single computational cell to define the entire width of the normal river channel and the local erosion is defined by the fluid shear stress that is averaged across the river cross-section. Because of the non-uniform nature of the actual erosion, and the fact that the model uses an average shear stress to define the erosion, it is highly probable that river bed erosion occurs long before the model predicts its occurrence. Furthermore, the PCB concentration of the sediment that is eroded in the model is the average of the local concentration over that cell. Thus the modeling uses cell averages that reduce the very broad PCB concentration distributions to a local average, and then redeposits that PCB-laden sediment at this average concentration. The net effect of the modeling therefore must be to smooth out the distribution of PCBs, but in reality this has not occurred in the significant length of time in which, according to the model, it should have occurred.

Response 3A-JL-2:

See Appendices B.1 and B.6 to the FMD. In addition, cross-channel differences in erosion and deposition patterns were relatively consistent for 21 of 53 profiles examined. Where cross-channel differences do occur, in many cases (20 of 53 profiles), these differences are so extensive that process aggregation and averaging would be required even if the river were represented as three (or more) grid cells across the channel.

The model framework does not act to “smooth out” PCB concentration gradients in river sediments (or floodplain soils) over time. As shown in the results of the calibration, validation, and example simulations, cell by cell differences in surface sediment PCB concentrations occur over the course of the simulations.

3.2 GENERAL TOPIC 3B: FLOW/VELOCITY/SHEAR STRESS

3.2.1 Comment Summary

Reviewer Name	Page Number/Line Number
Adams	Page 6, lines 26 to 43
Bohlen	Page 5, lines 20 to 37
	Page 9, line 29 to page 10, line 2
	Page 14, lines 20 to 25
Garcia	Page 5, lines 4 to 9
List	Page 1, lines 29 to 34
	Page 4, lines 27 to 41

3.2.2 General Topic Summary

Water in a river channel or on the land surface moves (flows) by the action of gravity. In the Housatonic River, the flow of water through the river channel or across the floodplain can transport and redistribute bed sediment, floodplain soil, and associated PCBs. As water flows, it is subject to resistance (drag) along the boundaries (surfaces) of all material it passes. The balance between gravity and drag forces along the flow path determines the velocity and depth of the flow. The force that flowing water exerts on the river bed and bank, as well as on floodplain when the water depth exceeds the height of the river banks, is described in terms of the shear stress. The hydrodynamic model in EFDC is used to simulate flow, depth, velocity, and shear stress throughout the Housatonic River system and floodplain. The velocity and shear stress results are used in subsequent EFDC sediment transport and PCB transport and fate computations.

Four Reviewers commented on the ability of the model to accurately simulate flows and velocities, which in turn affect bed shear and the resuspension of bed sediment into the water column. The Reviewers, however, were divided on their evaluations of this aspect of the modeling study. Two Reviewers indicated that the flow simulations were either “good” or “excellent,” but a third Reviewer felt that flow was not properly simulated due to errors in the timing (vs. magnitude) of flows. One Reviewer who was satisfied with the flow simulation specifically noted the timing issue, but indicated that such discrepancies were “not important.”

The fourth Reviewer believed that the size of the model grid cells precluded evaluation of processes, such as flow, that vary on sub-grid scales.

Two Reviewers generally had less confidence in the resultant bed shear and, as a consequence, the ability of the model to accurately simulate suspended solids concentrations. Both of these Reviewers felt that there were insufficient TSS data to make a satisfactory evaluation, but one concluded as a result that the model was unlikely to provide adequate prediction of PCB fate and transport.

3.2.3 General Topic Response

The Reviewer comments on this topic are diverse; therefore, no general response is provided. Responses to specific comments are provided below.

3.2.4 Response to Specific Comments on General Topic 3B

Eric Adams

Flows and Velocities

Flows were simulated for the 10.5 year Phase 2 calibration period (Figures 4.2-4 through 4.2-14). Unfortunately there is not much data to compare with. The most data are for early 1999 when the model consistently under-predicts flow. (Yet it seems to do well in the later validation periods, based on the pressure transducer data.) Similar agreement with stage suggests that average velocities should be pretty good. (Figure 4.2-26 shows reasonable agreement between measured and predicted velocities, but the predictions are generally too low; EPA argues that discrepancies are due in part to coarse grid resolution.) During storms there was reasonable agreement, with moderate overprediction of flow at low flow. Agreement during storms shows times of over- and under-prediction, but they seem OK on average. The model seems to do reasonably well simulating the extent of overbank flow during August 1990 (Figure 4.2-25). Model statistics seem generally within ranges specified by the QAPP and there are few indications of bias.

In the validation period, flows and velocities look similar to or better than the calibration period. There are some errors in the timing of flows, but these are not important. No major bias is seen.

Response 3B-EA-1:

EPA agrees with the Reviewer that the model framework hydrodynamic results are quite good with little bias.

W. Frank Bohlen

This matter of flow phase and velocity and associated effects on transport could be better evaluated if the report had provided a more complete discussion of measured vs. modeled velocities when presenting the results of EFDC hydrodynamic calculations (see pp.6-31 e.g.). Although the model uses HSPF generated flows at the upstream boundary, the resulting simulations are to some extent independent of the HSPF generated flow/stage through PSA and as a result may be less sensitive to this matter of stage timing. Speeds were measured using both electromagnetic and acoustic doppler currents meters for short periods of time at several locations within the PSA (see Figs 4.2-26 and 6.2-8 (attached)). These comparisons, while taken to be generally acceptable, indicate to me that there is very likely a substantial difference in measured versus modeled sediment transport associated with these differences in velocity. These flow induced differences will require substantial “calibration” within the model to yield reasonable estimates of sediment/contaminant flux. Given the non-linearity of both the velocity/shear stress and the shear stress/transport relationships it will be unlikely that such calibration will result in accurate simulations across a wide range of flows. This may be the principal reason that the model, at least some locations, seems to over-predict TSS values at low flows while under-predicting them at high flows. The significance of such variations will tend to increase with the duration of the model run and may become more of a problem during the extended runs planned for the Corrective Measures Study.

Response 3B-FB-1:

The small differences in the timing of simulated and measured flows are not significant, as noted by other Reviewers (see comment above from Eric Adams). Much of the difference is attributable to the timing of flow estimates computed using HSPF and used to drive EFDC. EPA believes that the hydrodynamics are appropriate for the goal of the modeling study. The differences in simulated and measured TSS concentrations are largely attributable to uncertainty in model upstream boundary conditions and tributary loads rather than small differences in hydrograph timing.

This depends to some extent on the characteristic being studied. For stream flow and stage model/data comparisons are based on a relatively long data set covering a wide range of seasonal, annual and intra-annual conditions at a number of sites throughout the PSA. The resulting comparisons are clearly sufficient to evaluate model capabilities over a relatively long period of time.

Moving to velocities and ultimately shear stress involves a significantly shorter data set at a limited number of locations. This does not necessarily mean that the data are inadequate since these characteristics are not expected to significantly change with time. As discussed above, however, it is the comparisons presented in the MVR that are less than sufficient. Better use of the available data would be the place to start. Careful review of the results of these analyses may then point to the need for additional data from differing locations and/or modified measurement procedures.

Response 3B-FB-2:

> EPA agrees with the Reviewer that the model-data comparisons presented are sufficient to evaluate model capabilities over a relatively long period of time. See FMD Sections 3 and 4 for presentation and discussion of model framework results.

This model system is not yet ready for use in the CMS. Several issues remain to be addressed. First, as discussed above, the hydrodynamics must be more thoroughly verified so as to insure accurate specification of boundary shear stresses. This factor will be central to future evaluations of selected remedial schemes such as capping and is presently essential within evaluations of sediment and PCB transport. The information and data provided in the validation report does little to build confidence in the present formulation.

Response 3B-FB-3:

> EPA disagrees with this Reviewer comment. The model framework can be used to evaluate the relative performance of potential remedial alternatives for the system. It should be noted that there is a fundamental misunderstanding between the many Reviewers and EPA regarding the use of the model framework for evaluation of alternatives. The framework will be used to examine the impact that potential remedial alternatives have on the system using a set of evaluation criteria established in the Consent Decree and Reissued RCRA Permit. The framework will not be used to engineer specific components of a remedy. This means that the framework is intended to examine only the large-scale, long-term impacts of remediation strategies.

Marcelo H. Garcia

As mentioned in the response to question one, the size of the computational grid makes it difficult (and meaningless) to compare model predictions for processes that take place at hydraulic and sedimentation scales determined by the flow rate and the main channel width. This is not the case for Woods Pond and the floodplains were the size of the computational grid is appropriate for the scale of the flow field.

Response 3B-MG-1:

> EPA disagrees with the Reviewer’s comment. The model framework has been calibrated and validated for individual events and longer periods. These comparisons demonstrate that the model framework simulated PCB transport in the Housatonic River system with the accuracy needed to achieve the modeling study goal. See FMD Sections 3 and 4.

E. John List

In general terms, the modeling tends to show excellent agreement for water mass fluxes, which is not at all surprising as models to simulate the flow of water in channels have been in use for a long time and a failure on this score would be cause for real alarm. There is only one adjustable parameter for such models and this is the friction factor for the stream, if this parameter is correctly chosen, and the stream cross-sectional area is correct, then the water levels and flow rates can be very accurately simulated.

Response 3B-JL-1:

EPA agrees with the Reviewer's comment. The results obtained demonstrate the proper functioning of the model framework.

The EFDC comparisons between measured and simulated flow and stage offered in Figures 6.2-3 of the Model Validation Report for years 1979-1999 present essentially no measured data and should not be offered up as part of the validation. The comparisons for years 2002-2004 are convincing that EFDC does a good job of representing flow and stage. However, the equivalent time series of Total Suspended Solids (TSS) concentrations for 2002-2004 (Figures 6.2-19 through 6.2-22) do not really provide an adequate basis to conclude that the sediment transport is being properly modeled (which is not surprising considering the points made above). The field data are simply too sparse to make any meaningful comparison between simulation and measured data. Even the data for two specific flow events modeled (Figures 6.2-23 and 6.2-24) offer only a few measured data points for comparison. The use of actual measured sediment rating curves either for calibration or for analysis is urged. The PCB sediment data offered in Figures 6.2-49, 6.2-50, and Figure 6.2-51 indicate just how difficult is the process of validating the PCB fate and transport model and how unlikely it will be that the existing model will provide any real hope of predicting the future fate and transport of the PCB.

Response 3B-JL-2:

Although the density of data for TSS and PCB concentrations in the water column is less than for flow and stage measurements, comparisons of model results to the site data demonstrate that the model successfully simulates sediment and PCB transport in the Housatonic River over the 26-year validation period.

Model performance is often driven by the solids and PCB loads entering the system from the upstream boundary. These upstream boundary conditions are rating curves where concentrations are parameterized as functions of flow, as described in FMD Appendix B.2. However, the uncertainty associated with these rating curves is large as a result of natural variations in solids and PCB concentrations. For example, solids and PCB concentrations can vary by more than a factor of 100 at any one flow rate. Even if rating curves for downstream locations could be developed, they would have very large uncertainty due to natural variability. Consequently, EPA believes that direct comparison of model results to individual data (however spare) is the most reliable means to evaluate model performance.

Further, EPA believes the model framework can successfully simulate long-term PCB transport and fate in the Housatonic River because the model calibration approach helps ensure that all processes are properly parameterized and also because the model meets the established performance targets. Evaluations of model performance are documented in FMD Section 3.

3.3 GENERAL TOPIC 3C: FLOODPLAIN INTERACTIONS

3.3.1 Comment Summary

Reviewer Name	Page Number/Line Number
Bohlen	Page 8, lines 24 to 37
	Page 14, lines 26 to 31
Endicott	Page 20, line 33 to page 21, line 3
Garcia	Page 1, lines 25 to 37
	Page 4, line 19 to page 5, line 2

3.3.2 General Topic Summary

In addition to aggradation and degradation of bed sediment, the banks of the Housatonic River are routinely modified by erosion, and interaction of the river channel with the floodplain occurs during periodic out-of-bank flood events. As riverbanks erode, solids and PCBs in the banks enter the river channel where they may be transported farther downstream. Similarly, periodic out-of-bank flow events transport solids and PCBs across the floodplain. Due to increased vegetative resistance relative to the river channel, flow velocities through inundated portions of the floodplain are expected to be smaller than velocities in the channel. Consequently, materials transported from the channel accumulate in the floodplain (Leopold et al., 1964; Knighton, 1998). In addition to vegetative resistance, the magnitude of accumulation of solids and PCBs on the floodplain may also depend on other factors, such as the frequency and duration of out-of-bank events; upstream sediment, PCB, and overland solids inputs; floodplain topography; and the deposition characteristics of the materials in transport (see Appendix B.7 to the FMD).

With regard to sedimentation and related processes in the floodplain, one Reviewer noted that the model accounts for bank erosion and floodplain sedimentation/resuspension but does not consider transport of contaminated leaf litter or wash-off of material adhering to floodplain vegetation following flood events. The Reviewer questioned whether any thought had been given to including these processes in the model and suggested that phytoremediation might be considered in the floodplain.

Another Reviewer underscored the importance of processes in the floodplain and listed them along with a wide variety of other processes that he felt needed to be better understood and improved in the simulation. He further noted that the model had been validated over a long time frame (i.e., 26 years), but the majority of transport occurs during relatively short storm events, and he recommended that the model also be validated over shorter periods of high flow. Finally, the Reviewer suggested that calculation of a "Floodplain Dimensionless Number" for specific reaches could be used as a tool to evaluate the validation.

3.3.3 General Topic Response

The Reviewer comments on this topic are diverse; therefore, no general response is provided. Responses to specific comments are provided below.

3.3.4 Response to Specific Comments on General Topic 3C

W. Frank Bohlen

The side bank issue affects both the margins of the river channel and the floodplain. All indications suggest that this latter area is primarily a sink for sediment and PCBs. As I understand it, the model treats each of the grids on the floodplain in a manner similar to those in the river channel and seeks to erode the soil surface by flow induced shear during flooded conditions. Given the presence of vegetation this very seldom occurs leading to continuing deposition over most of the area. What sediment and PCB that is supplied by the floodplain comes from aperiodic failure of the floodplain margin or sidebank.

> **Response 3C-FB-1:**
>
> Floodplain interactions are not exclusively limited to solids and PCB deposition. The model framework can simulate both erosion and deposition in the floodplain. Over the course of the simulations, some erosion may occur in the floodplain. This is just one example of why it is so important to actively include the floodplain in the model domain. Further, bank erosion is a significant source of both solids and PCB. Over the course of the 26-year model validation period, bank erosion amounts to approximately 13% of the solids mass budget and approximately 28% of the PCB mass budget. See FMD Section 4 for further detail.

If correct, this scheme seemingly neglects any transport associated with the movement of leaf litter and/or the rainfall induced wash-off of materials adhering to the surfaces of vegetation following flood inundation. Has consideration been given to the inclusion of these factors in the model? If not, is there a solid basis for their neglect? It may be that this is a subject that could be quantified in the revised monitoring program recommended above. It may also be that phytoremediation should be included in this program (if it has not previously been investigated) and/or included in the upcoming CMS.

Response 3C-FB-2:

EPA believes that the potential for PCB loading to the river via the mobilization of leaf litter or particle wash-off from vegetation is an extremely small component of the overall mass fate of PCBs in the Housatonic River system because the mass of particles or leaf litter involved is expected to be orders of magnitude less than the mass entering by bank erosion. The solids and PCB masses entering the river by bank erosion is, however, quantified in the model framework.

Identification and analysis of potential remedial alternatives will be performed by GE in the CMS. It should be noted, however, that phytoremediation is not a remediation technique that has been found to be effective for PCBs.

Next, additional work is required to develop an accurate formulation of sediment transport. The suggestions of Dr. Lick in terms of both equation parameters and the structure of the sediment bed should be carefully evaluated. There seems to be an abundance of data and experience to suggest that 6in is an overestimate of the active bed thickness. There is also concern that the model as it exists may be biased to high flow conditions. Add to these questions regarding side bank erosion and floodplain dynamics.

Response 3C-FB-3:

See Section 2.3.2 and Appendix B.8 of the FMD for a review of sediment transport process formulations and bed representation. Active bed thickness in EFDC is on the order of 4 to 7 cm (1.5 to 2.75 inches). Further, the thickness of the active layer in the bed has been verified through using sediment profile imagery as presented in Appendix A.3 to the FMD. Additional discussion of the representation of the sediment bed is presented in FMD Appendix B.5.

Douglas Endicott

As pointed out by GE, predictions of PCB concentrations in flood plain soils are untested. There appear to be substantial increases in PCB concentrations in the top 6” of much of the floodplain soil in both example projections, which is somewhat surprising. However, since no data have been shown to establish rates of PCB concentration change in floodplain soils, the “reasonableness” of the flood plain projections cannot be determined. Really, this illustrates the absurdity of devoting so many EFDC model grid cells to the flood plain, which it seems dictated using a single footpath of tiles for the grid of the river.

Response 3C-DE-1

EPA disagrees with the Reviewer’s opinion regarding the “absurdity of devoting so many EFDC model grid cells to the floodplain…” Because 60% of the PCBs within Reaches 5 and 6 (which contain an estimated 90% of the PCB mass) are estimated to be in the floodplain, this portion of the domain deserves the attention provided by the grid. This is required to obtain an understanding of the response of the potential remedial alternatives with regard to the commensurate reduction in risk.

In response to comments in General Topic 3 and General Topic 4, EPA indicated that trading grid resolution on the floodplain for resolution in the channel would not achieve substantial benefits in grid implementation because calculations are performed in floodplain cells during a small portion of the simulation (when these cells are wet). Conversely, decreasing the number of cells on the floodplain and increasing the cells in the channel would dramatically increase run time.

In response to the Reviewer's assessment of "substantial increases in PCB concentrations in the top 6" of much of the floodplain soil," it is noted that these represent large percentage increases from initial concentrations (e.g., a 100% change if concentrations increased from 0.5 to 1 mg/kg). Uncertainty in establishing initial conditions is a major factor that contributed to increases in floodplain soil concentrations simulated by the model. Both of these factors are discussed in more detail in Appendix B.3 of the FMD.

Marcelo H. Garcia

While a substantial effort has gone into improving the capabilities of the Housatonic River Model since its calibration, it is clear that there is still a long way to go before the model can truly be used as a "predictive tool" to quantify future spatial and temporal distributions of PCBs (both dissolved and particulate forms) within the water column and the bed sediment.

This Reviewer is of the opinion that the Housatonic River Model will need to be continuously improved until all the processes (biological and physical bed sediment mixing, streambank erosion, floodplain deposition, etc.) relevant to the transport and fate of PCBs are not only accounted for in the model but are also well represented and based on sound knowledge of sediment transport mechanics, stream biology, ecology and morphodynamics.

Response 3C-MG-1:

EPA disagrees with this comment because the ability of the model framework is already well-established (as documented in Sections 3 and 4 of the FMD, which summarize previous model performance). With regard to bank erosion and floodplain interactions, these processes are included in the model, as described in FMD Appendix B.7. All relevant processes are included, at the relevant scales, and are based on the current understanding of these processes.

Regarding floodplain sedimentation, the model has been used to predict process-based (advection, erosion, deposition, volatilization, etc) sediment and PCB fluxes (Kg/year) for the main channel and the floodplain, over the validation period (Figures 6.2-62 and 6.2-63). The analysis results in an estimate of yearly fluxes. However, it is well known that sediment erosion and transport is most prominent during flood events, which can have duration of a few hours to several days. Thus it would seem that the model should be validated for time scales that are relevant to the processes involved.

Since a large percentage of the river system consists of floodplain, it is important to ensure that the model can indeed capture the process of floodplain sediment/PCB deposition. To this end, a simple one-dimensional approach was suggested that could be used to estimate a "Floodplain Dimensionless Number" for different reaches of the Housatonic River (Garcia, 2006-technical note on floodplain sedimentation). Once such number is calibrated for each reach, it will be possible to test if indeed the model can predict floodplain depositional rates that result in similar Floodplain numbers for different flooding conditions (hydrologic event scale), thus helping in the overall validation of the model. This approach will clearly show the capabilities of the model. If there is any hope of predicting PCBs fate in the floodplain, sedimentation has to be properly simulated by the model.

Response 3C-MG-2:

See Appendix B.7 to the FMD for specific discussion of the "Floodplain Number." Although the duration of flood events is typically short, it should be noted that the storm event water column sampling program was specifically designed to provide measurements of solids and PCB concentrations during events. Focusing on Reaches 5 and 6, all inputs and outputs to the model domain during events can be reasonably determined from the event data in combination with other flow data. From this information, the mass of solids and PCBs accumulating in the sediment bed and floodplain areas can be successfully determined from the difference between inputs and outputs. This is a reasonable approach because it is based on a fundamental engineering principle, conservation of mass.

Furthermore, application of a dimensionless Floodplain Number (FN) is fundamentally flawed because the specific types of measurements needed to develop FN values do not exist for the Housatonic River system. In the absence of measurements, the only values that could be used to compute a FN would themselves be computed by averaging EFDC results. Consequently, a difference between an inferred FN value and full-scale EFDC results would indicate only a difference between the one-dimensional assumptions of the scaling analysis used to derive the FN and the two-dimensional conditions simulated by EFDC during out-of-bank events. More importantly, if values for parameters needed to compute a FN were known or could be determined from measurements, these values could be directly compared to values computed in EFDC without the need for an approximation using the FN. As a result, the FN approach is not informative for evaluating model performance.

3.4 GENERAL TOPIC 3D: BED LAYERING AND BED MIXING

3.4.1 Comment Summary

Reviewer Name	Page Number/Line Number
Adams	Page 2, line 29 to page 3, line 23
Bohlen	Page 10, lines 25 to 32
	Page 13, lines 5 to 11
	Page 14, lines 26 to 31
Endicott	Page 3, lines 13 to 27
Garcia	Page 6, lines 4 to 7
Lick	Page 6, line 9 to page 7, line 43

3.4.2 General Topic Summary

The specification of vertical layers (compartments) in the sediment, and the extent to which material communicates between different layers, controls simulated sediment and water contaminant concentrations that drive food chain exposures. One element of sediment column compartmentalization of particular importance to PCB transport and bioaccumulation model development is the delineation of the active layer within the sediment bed. The depth (thickness) of the active layer can be a controlling factor when estimating food chain exposure concentrations.

The term “active layer” may be one of the most confusing phrases used in PCB transport and bioaccumulation studies because it has distinctly different meanings, depending on context. The commonality in all of these contexts is that the active layer is ultimately expressed as a depth into the sediment column over which mass transport processes or exposure can occur. Consequently, the compartmentalization and parameterization of layers in the sediment column is an important consideration.

Several Reviewers commented on the sediment bed layering in the model, particularly with regard to the specification of the thickness of the well-mixed surface layer and the rate of mixing between layers. In discussing this issue, all of these Reviewers focused on mixing attributable to benthic infauna and epifauna, and all agreed it is an important issue because it directly affects the rate of change in sediment PCB concentrations as a result of simulated remedial alternatives. Reviewers also agreed that the specified depth of the well-mixed layer was too thick, although there was apparently some confusion about what depth(s) were used in the Validation.

3.4.3 General Topic Response

The mixing of particles and sorbed contaminants in surface layers of the sediment bed is a complex, site-specific, time-dependent process. The extent and rate of particle mixing in the sediment bed can vary widely. The causative agents of sediment mixing also vary from site-to-site and time-to-time. For example, mixing can be caused by benthic organisms and macrofauna (bioturbation), physical causes such as turbulent burst and sweep, and chemical processes such as gas ebullition.

Throughout the Peer Review, numerous opinions were offered with respect to the extent and rates of particle mixing. Most Reviewers focused on bioturbation attributable to benthic organisms, largely neglecting mixing by other factors. Comments made on behalf of GE furthered this discussion and also focused exclusively on bioturbation by benthic organisms (QEA, 2006). In contrast, while the extent and rates of sediment mixing in the model framework were based on literature review and site-specific determinations based on measured benthic organism types and densities, this was done with the explicit recognition and understanding that other factors may control mixing at any site over time.

Further review of mixing by benthic bioturbation was performed and is summarized in Appendix B.4 to the FMD. Sediment profile imaging (SPI) was also conducted and is summarized in FMD Appendix A.3. In the SPI study, a high-resolution camera was inserted into the sediment bed of the river to take detailed photographic images of sediment profiles at 52 stations in Reaches 5 and 6. A total of 175 sediment profiles were collected. Based on the interpretation of these profiles, average sediment mixed layer depths were determined to range from 3 to 16 cm. Furthermore, in many locations rapid physical processes were determined to be the primary agent for mixing. The model formulation is consistent with these observations, specifically with respect to the parameterization of surface layer thickness. Moreover, these results suggest that it might be reasonable to specify faster rates of mixing in the model.

3.4.4 Response to Specific Comments on General Topic 3D

Eric Adams

Biomixing and Bioavailable Depth

Originally, the model used a bioavailable depth of 15 cm over which the biomixing coefficient (bio-diffusivity) was $D_b=10^{-9}$ m^2/s, or approximately 1 cm^2/d. I commented that this seemed quite large, and EPA agreed. They have changed their formulation to utilize a subduction velocity, V_s, which they get from a literature review of the rates of sediment mass reworking per organism per time, and site-specific data on organism abundances. Their chosen velocities are approximately $V_s \sim 10^{-9}$ m/s over the top 4-7 cm, and $\sim 10^{-10}$ m/s over the next 6-8 cm. In the absence of direct measurements of mixing, this may be the best approach, but it is noted from figures presented at the May 10 Document Review Meeting (DRM) that there is tremendous variability in organism reworking rates, suggesting much uncertainty.

Assuming a vertical distance of 5 cm between the top two sediment layers, their equivalent new values of bio-diffusivity are $\sim(10^{-10}$ to 10^{-9}m/s)(0.05m) $\sim 5\text{x}10^{-12}$ to $5\text{x}10^{-11}$ m^2/s or ~20 to 200 times smaller than previous. These are probably more reasonable, but it is difficult to assess whether or not they are right, because there is not much vertical variation in the existing sediment PCB concentrations. However, following remediation, the vertical concentration gradients could increase substantially, so this is an important process.

This also raises the issue of vertical resolution. If, following remediation, clean sediment is overlain by a thin layer of contaminated sediment, the numerical model will immediately mix the contaminant over the top layer (say 5 cm), whereas, with a bio-diffusivity of $5\text{x}10^{-7}$ cm^2/s, mixing will take $(5cm)^2/5\text{x}10^{-7}$ cm^2/s or about 1.5 years to achieve this mixing. Thus, from a numerical modeling standpoint, **the vertical grid size is too large.**

The sediment-water interface can rise or fall due to deposition and erosion, and EPA argues that this should not affect the rates of mixing. This may not be true since you could expect different rates of organism mixing in relatively fine grained sediments that have been recently deposited, versus older, more consolidated sediments that have been eroding. Indeed, Figure 13 of the DRM handouts on sediment mixing shows a strong increase in V_s with percent fines. However, to honor their assumption, they make V_s dependent only on the depth below the (moving) interface. But because the individual layer thicknesses are changing, the amount of vertical mixing will change. (A constant value of V_s will result in more mixing between thick layers having the same concentrations as thin layers.) Using a bio-diffusivity (dimensions of L^2/T) would take this effect into consideration since it is effectively V_s times the mixing length.

Response 3D-EA-1:

See the General Response above. The Reviewer comment focuses exclusively on biological mixing. Biological mixing is believed to be the slowest process of all the processes that are believed to cause particles in upper sediment layers to mix over time. As presented in Appendix A.3 to the FMD, sediment profile images conclusively show that mixing depths in the Housatonic River range from 3 to 16 cm. The model framework is consistent with these site-specific observations.

W. Frank Bohlen

The model/data comparisons for PCBs should be extended to include consideration of distributions over the vertical. Any assumption of a well mixed layer extending over depths in excess of a few inches doesn't seem to agree with field data (see e.g. Fig. 4-21c, BBL,2003, attached). How are these differences to be reconciled? (i.e. use of 6" well mixed layer vs. detailed core data sowing little mixing beyond 1-2in). These core data and associated radio-dating also allow estimates of sedimentation rate to be compared to the mass flux data provided by the model. This comparison was not part of the validation report. It would provide an additional check on model results and is recommended.

Response 3D-FB-1:

See FMD Section 2.3.2 and Appendix B.8 for a review of sediment transport process formulations and bed representation. Active bed thickness in the model (specifically EFDC) is on the order of 4 to 7 cm (1.5 to 2.75 inches). Further, the thickness of disturbed sediment in the bed has been quantified through sediment profile imagery as presented in Appendix A.3 to the FMD. Additional discussion of the representation of the sediment bed (with supporting sediment profile imagery) is presented in FMD Appendix B.5. See General Response above.

In addition to testing of the erosion formulation additional sensitivity analyses should be performed to assess model response to the thickness of the active sediment column. This has been a matter of concern for some time. Experience in other riverine/bayou systems as well as the detailed core data (Fig. 4-21c, attached) indicate to me that the active bed used in the model is too thick. It may be that this specification in terms of physical transport characteristics has relatively little effect on overall model results (although that might bring up another set of questions). A test of model response to this characterization is recommended.

Response 3D-FB-2:

With respect to active layer thickness, see General Response above. Although the thickness of the active sediment layer was not explored as part of the model sensitivity analyses, it was included in the model uncertainty assessment conducted in the MVR.

Next, additional work is required to develop an accurate formulation of sediment transport. The suggestions of Dr. Lick in terms of both equation parameters and the structure of the sediment bed should be carefully evaluated. There seems to be an abundance of data and experience to suggest that 6in is an overestimate of the active bed thickness. There is also concern that the model as it exists may be biased to high flow conditions. Add to these questions regarding side bank erosion and floodplain dynamics.

Response 3D-FB-3:

See Response to 3D-FB-1. Bank erosion is constrained to reproduce site-specific long-term and short-term average erosion rates. Also see the General Response above and FMD Appendix B.7.

Douglas Endicott

The vertical representation of the sediment bed in the PCB fate and transport model: The thickness of the surficial sediment layers have been specified as either 4 cm (Reach 5A) or 7 cm (Reaches 5B thru 6). These layers are assumed to be well-mixed, via a combination of physical and biological processes. In addition, sediments from the surficial layers are mixed with deeper layers, down to a depth of 30 cm below the sediment-water interface. Analysis of benthic invertebrate data is offered to justify these assumptions. However, these data show that below Reach 5A the benthos are predominantly filter feeders or surficial sediment feeders/dwellers. Such organisms are unlikely to cause sediment mixing below the top couple of centimeters. 7 cm of surface sediment mixing is too deep, and mixing of sediment between the surface layer and deeper layers is inconsistent with the life history data for the resident benthos in these reaches. The consequence of too much sediment mixing (and too-thick surficial layers) is that the PCB fate and transport model will predict an unreasonably slow rate of decline in surficial sediment PCB concentrations, which is observed in both Phase 2 calibration and validation periods.

Response 3D-DE-1:

EPA strongly disagrees with this Reviewer comment. First, it must be noted that this comment is predicated on an erroneous assumption that surface sediment PCB concentrations are declining rapidly. As documented in the RFI, FMD Appendix A.1, and elsewhere (see General Topic 1 and General Topic 2), the rate of change of PCB concentrations in surface sediments is slow and cannot be distinguished from zero (i.e., slow to no change over time). Second, although the Reviewer’s comment recognizes that physical mixing processes can occur, it presumes that benthic organisms are the sole cause of mixing in sediments. In contrast, SPI results demonstrate that well-mixed layer thicknesses in Housatonic River sediment range from 3 to 16 cm on average. In many of these profiles, the cause of mixing was attributed to physical and chemical processes rather than biological processes.

Marcelo H. Garcia

In the case of low flow conditions, the model predictions seem to be very sensitive to diffusion parameters, suggesting the representation of pore water diffusion as well as the thickness of the so-called "mixed layer' need to be carefully analyzed.

Response 3D-MG-1:

See the General Response above and FMD Appendix A.3. The representation of the mixed layer in the model is consistent with site conditions as confirmed through sediment profile imagery.

Wilbert Lick

The Sediment-Water Flux of HOCs Due to "Diffusion"

The non-erosion/deposition flux of contaminants from the sediments to the overlying water is primarily due to molecular diffusion, bioturbation, and ground-water flow. Each of these processes behaves in a different way and hence needs to be modeled in a different way. EPA has chosen to describe all of these processes by means of a "diffusion" model based on the concept of bioturbation and the assumption of a well-mixed layer. This is the conventional, but not necessarily accurate, approach. It is not accurate simply because the mass transfer approximation (which is not a diffusion approximation) actually used by EPA does not describe or adequately approximate the HOC fluxes of molecular diffusion, bioturbation, or ground-water flow, not even in functional form. The correct functional form (especially its dependence on time) is important because, otherwise, even calibration doesn't work for long term predictions.

EPA did an extensive review of the literature on bioturbation and listed 139 documents of which 43 were retained for detailed review. This listing is somewhat misleading. Of the 43 most relevant documents, almost all are general observations, surveys of the literature, or even surveys of surveys; only about six report quantitative data or laboratory measurements of mixing due to benthic organisms. For example, the figure shown at the last meeting entitled "Bioturbation and Bioavailable Sediment Depths" is from Clarke et al. (2001) and is their interpretation of what organisms do. There is no data (given by Clarke et al. or anywhere else) to support this figure. Clarke et al. is an excellent manuscript, but it is another survey. There is no new data there. None of the documents listed by EPA report on the sediment-water flux or sediment mixing of hydrophobic organic chemicals (HOCs) due to benthic organisms. Since then, EPA has listed additional reports concerned with the flux of HOCs due to benthic organisms. However, these HOCs had relatively low partition coefficients. The only quantitative data that I know of on the effects of benthic organisms on the flux of HOCs with large partition coefficients and their resulting vertical distribution in the sediments is that by Luo et al. (2006) which is attached. Some of my comments are based on this article.

In EPA's modeling, assumptions are (1) a constant (independent of space and time) sediment-water mass transfer coefficient, k, with a value of 1.5 cm/day and (2) a surficial, well-mixed layer whose thickness is constant in time but varies spatially from 4 cm upstream to 7 cm downstream. Measured biomass varies by about a factor of 20 from upstream to downstream.

Benthic mixing is described by EPA in terms of a mixing rate (a diffusion process), also termed a subduction velocity (a convection process), a quantity which I believe is used in the model as a mass transfer coefficient between the sub-surface sediment layers. For the biologically mixed layer, the subduction velocity (values from EPA's table) varies from about 1 x 10^{-9} m/s (1 x 10^{-2} cm/day) upstream to 2 x 10^{-9} m/s (2 x 10^{-2} cm/day) downstream. This factor of two between upstream and downstream seems surprising since the biomass increases by a factor of 20 in the downstream direction. Even more surprising is that the mass transfer coefficient, k, is assumed constant everywhere at 1.5 cm/day. Why doesn't k increase downstream as the biomass increases by a factor of 20?

A surficial well-mixed layer whose thickness is constant in time is assumed in the analysis. An approximate and minimum time for formation of this layer can be calculated from $t = h/v_b$, where h is the thickness of the layer and v_b is the subduction velocity. For the upstream area, h = 4 cm, v_b = 1 x 10^{-9} m/s = 1 x 10^{-2} cm/day, and therefore t = 400 days. For the downstream area, h = 7 cm, v_b = 2 x 10^{-9} m/s = 2 x 10^{-2} cm/day and therefore t = 350 days. In other words, these so-called well-mixed layers are not formed instantaneously and take a minimum of 350 to 400 days to form.

This becomes a little confusing upon examination of the figure presented at the meeting entitled "Contribution to Db from different groups of benthos". Upstream, Db (for all benthos) is approximately 2 x 10^{-3} cm/day while downstream, oligochaetes (the main vertical burrowers and subductionists) contribute a Db of approximately 2.5 x 10^{-2} cm/day. The upstream number is an order of magnitude less than the number in the table cited above as v_b, is probably correct, but gives a time for formation of the well-mixed layer of almost 2000 days (6 years).

Despite the confusion, the numbers for Db are probably correct (to within less than an order of magnitude). This demonstrates (as does Luo et al. more accurately and convincingly) that so-called well-mixed layers for HOCs, if they exist, take a long time (years) to form. Is there any evidence that a well-mixed layer even exists in the Housatonic? I don't believe so.

The mass transfer coefficient, k, for the transport of HOCs by molecular diffusion alone is approximately 1.2 cm/day and decreases slowly with time at a rate which decreases as K_p increases (Deane et al. 1999, Lick et al. 2006, attached). If there is only a small number of organisms, EPA's value of k = 1.5 cm/day compares well with this number. However, with benthic organisms present, Luo et al. give a mass transfer coefficient for HOCs that varies up to 10 cm/day (for benthic organism densities of $10^4/m^2$) and somewhat higher for very dense concentrations of organisms; these values for k are much higher than those that EPA assumes.

No consideration is given to ground-water flow, which can be significant, is a convection and not a diffusion process, and does not involve a well-mixed layer of any sort.

Response 3D-WL-1:

EPA disagrees with the Reviewer's comment. It should be recognized that this comment is predicated on the assumptions that: (1) particle mixing must always be very slow as controlled by benthic bioturbation rates; and (2) a well-mixed layer does not exist and will not form. Specifically, the Reviewer asks if there is any evidence that a well-mixed layer exists in the Housatonic River. As documented in Appendix A.3 to the FMD, 175 sediment profiles images were obtained at 52 stations to further document the knowledge of sediment mixing in the Housatonic River. Based on these observations, well-mixed sediment layers exist and range in thickness from 3 to 16 cm. In many of these profiles, the cause of mixing was attributed to physical and chemical processes rather than biological processes. Therefore, in specific answer to the Reviewer's speculations, the sediment profile images conclusively demonstrate that a well-mixed layer exists in the Housatonic and that apparent rates of mixing can be much larger than inferred by the Reviewer.

Further is should be noted that the "diffusion" term on which the Reviewer comments is the sediment-water mass transfer process. This mass transfer process represents an aggregation of individual processes that may include diffusion, groundwater advection, bioirrigation, and others. The important issue for a practical model application is whether the mass transfer of PCBs from the bed to the water column is adequately represented. At the scales of interest for the modeling study, there is no need to further determine which of the underlying mechanisms are actually transporting PCBs from the bed as long as the net flux from the bed is accurately represented. In the specific case of the Housatonic River, the sediment-water mass transfer of PCBs from the bed is one of the most well-defined elements of the model framework because the scale of this flux was determined directly from measurements.

See Appendix B.10 to the FMD for further details regarding the sediment-water mass transfer of PCBs.

4. GENERAL TOPIC 4: MODEL EFFICIENCY, COMPUTER RESOURCES, AND RUN TIME

4.1 COMMENT SUMMARY

Reviewer Name	Page Number/Line Number
Adams	Page 12, lines 37 to 44
	Page 13, line 21 to page 14, line 37
Bohlen	Page 14, line 32 to page 15, line 15
	Page 15, lines 16 to 23
Endicott	Page 7, lines 20 to 25
	Page 8, lines 8 to 12
Garcia	Page 2, lines 33 to 37
	Page 6, lines 13 to 17
Gobas	Page 4, lines 17 to 31
	Page 4, lines 38 to 45
	Page 14, lines 11 to 13
Lick	Page 10, lines 14 to 34
	Page 13, lines 17 to 19

4.2 GENERAL TOPIC SUMMARY

Mathematical models provide simplified or approximate representations of reality. In reality, chemical and physical properties can vary continuously in space and time. In a mathematical model, these continuous variations are represented by step-changes in discrete intervals of space (a grid cell) and time (the timestep). The use of a grid to represent a physical domain necessitates approximations of spatial variations in physical and chemical properties, because within each grid cell, a single representative value is used, even if properties within the grid cell vary in space. The continuous temporal change in a physical or chemical property is approximated by a sequence of steps, in which the property (e.g., concentration) is represented by a single value over the timestep. The finer the resolution in space and time, the more closely the step-changes in space and time can approximate the true continuous variations. (Temporal and spatial coverage of data, however, can impose limitations on the approximation of the continuous solution). The choices of spatial and temporal resolution have direct impacts on the run time and computer resources required for a particular simulation.

A number of techniques are used in model development to balance the conflicting objectives of (1) high temporal and spatial resolution, and (2) acceptable run times. These techniques include dynamic time-stepping, split time-stepping, and model linkage options. Dynamic time-stepping refers to changing the timestep during the course of the simulation, in response to changes in flow conditions. Split-time-stepping refers to the use of different timesteps in different parts of the grid, such as in the sediment bed, where changes are much slower than in the water column, or by skipping calculations on the floodplain, when cells are not inundated. EFDC includes options to use both dynamic and split time-stepping. Another technique involves the linkage between hydrodynamic, sediment transport, and contaminant fate submodels. EFDC runs all three of these modules simultaneously.

Some model codes perform the calculations for hydrodynamics and sediment transport together and then write output files, which are read as input to the contaminant fate model. It is also possible with some models to run hydrodynamics and sediment transport independently. Decoupling of models involves approximating more-finely resolved temporal variations in model results, such as flow or water depth, by an average value over some specified averaging period, for example, 1 hour. The choice of the averaging period represents a trade-off between the accuracy of the approximation of the averages to the finer time-scale computations versus file sizes required to pass hydrodynamic output to the sediment transport model or hydrodynamic and sediment transport output to the contaminant fate model. The effort to decouple an existing model depends on the structure of the code. Decoupling code includes re-structuring the code, developing averaging algorithms for each model output passed to subsequent submodel, and testing the results for mass balances and consistency with the results of the original fully integrated code.

Five of the seven Reviewers commented on the computer resources required to run the model and the consequent long run times necessary for the simulation of the effectiveness of remedial alternatives. A number of recommendations were provided for decreasing the run time, with four of the Reviewers suggesting the decoupling of various model components. Three of the

Reviewers recommended that some parts of the model could be used to generate “synthetic” sequences such as 1 year of hydrology that is repeated many times or approximate low-flow solutions that are substituted for long periods of low flow, interspersed with full solutions for high-flow events. Three Reviewers recommended that the model grid be simplified, in most cases by increasing cell size in the floodplain, and even, as recommended by one Reviewer, eliminating the use of the model in the floodplain and substituting a reach-specific dimensionless number that could be used to quantify PCB deposition.

4.3 GENERAL TOPIC RESPONSE

The physical characteristics of the Housatonic River present a challenging setting for grid development and contaminant fate and transport modeling, which is further complicated by the need to include the floodplain in the computational domain. The floodplain was included in the computational grid because it has been estimated that 60% of the PCBs present in the PSA are in the floodplain, and risks to human health and the environment were identified from exposure to PCBs in the floodplain. Coupling grid cells in the meandering channel with grid cells on the floodplain complicated the development of the grid, and required careful balancing of grid resolution versus simulation run time objectives.

The modeling team clearly recognizes the significance of run time issues associated with decadal-scale EFDC simulations, and appreciates the Reviewers’ thoughtful recommendations of potential options for reducing run time. Many of these same recommendations were made at the time of the Model Calibration Peer Review. These recommendations were considered then, and reevaluated based on the additional comments received following the Model Validation Peer Review.

Several Reviewers commented that decoupling the channel grid from the floodplain grid could provide a reduction in run time. However, EFDC already contains logic tests, which provide a means of bypassing computations for floodplain cells that are not inundated, so the floodplain and channel are effectively decoupled during the majority of the time when in-bank flows are simulated.

Reviewers suggested that increased grid resolution in the channel could be included without increasing run time, if there were a reduction in the grid resolution on the floodplain. The change in model run time resulting from decreasing grid cells on the floodplain and increasing grid cells in the channel goes beyond comparing the total number of grid cells in the model domain. The logic checks in EFDC, which allow computations on the floodplain to be bypassed if the cells are not flooded, results in calculations in floodplain cells at only a fraction of the time calculations are performed in channel cells. Including a higher resolution grid in the channel and decreasing the resolution on the floodplain would result in an increased number of cells that would be constantly active, and therefore would result in substantially increased model run time. These factors are discussed in Appendix B.1 of the FMD, and in Response to General Topic 3.

Four of the Reviewers recommended decoupling the hydrodynamic, sediment transport, and contaminant fate submodels. This concept was thoroughly explored by the modeling team during the Phase 1 Calibration (prior to the Calibration Peer Review). While conceptually straightforward, as one Reviewer pointed out, “the devil is in the details.” This task would have involved rewriting major portions of the code, and the development of averaging schemes to write files containing average transport information from one submodel, which could be read as input in the next submodel. Averaging would be required because saving output at every time step is not feasible, even with today’s large-capacity hard drives.

Members of the modeling team have participated in a similar effort to link decoupled models (a hydrodynamic and sediment transport model, ECOMSED, with the water quality model, RCA). This experience demonstrated that the task could be accomplished. However, it highlighted the fact that the undertaking introduces substantial uncertainties associated with estimates of the schedule for completion of the programming and testing and associated costs. The modeling team estimated that the impact to decouple the EFDC code would be to dramatically influence the Rest of River project schedule/budget and also introduce a high degree of uncertainty into the schedule/budget. It should be noted that this approach would have potentially benefited the calibration and validation phases of the modeling study to an even greater degree than the Corrective Measures Study (CMS), so the impetus was there to do it; however, the implementation and resulting uncertainties rendered it impractical for a Peer-Reviewed evaluation of model performance.

This topic was revisited based on the Reviewers' recommendations made as part of the Calibration Peer Review and Validation Peer Review. The modeling team reached the same conclusion, judging that the uncertainties/risks are of greater consequence than any issues associated with proceeding with the proven, linked model code.

The decision not to decouple EFDC was made with consideration of the simulations needed to support the CMS. It was recognized that some of these simulations may be considerably longer than the validation simulation of 26 years. It should be noted that the modeling during the CMS can utilize the capability of EFDC to include multiple PCB state variables in a single simulation. Scenarios that can be represented by the same hydrodynamic and sediment transport conditions (e.g., different estimates of residual PCB concentrations) can be simulated in a single simulation, providing an "economy-of-scale" to the evaluation. For different scenarios that include different sediment bed conditions, separate hydrodynamic and sediment transport simulations would be required to account for differences in bed roughness and grain size distributions associated with the differing representations of the sediment bed. In these cases, decoupling the individual sub-models of EFDC would not provide additional computational benefit.

Recent hardware advances have reduced model simulation times to between half and two-thirds of the run times experienced during the Phase 2 Calibration/Validation modeling effort. In addition, these faster chips are available in dual-CPU servers at a nominal cost. With these computers, a single computer can process two 50-year simulations in approximately 2 weeks. At a cost of $50,000 (the investment made by EPA in computing hardware in 2004 to conduct the modeling study), 20 50-year simulations could be completed every 2 weeks for Reaches 5 and 6.

The application of the model in the CMS will be discussed in the Model Working Group and subsequently proposed by GE and reviewed by EPA.

4.4 RESPONSE TO SPECIFIC COMMENTS ON GENERAL TOPIC 4

Eric Adams

The issue of model resolution conflicts directly with the issue of computational time. Clearly, with the current coupled in-channel/overbank modeling system, one cannot afford the extra computational cost of reducing Δy, and still be able to afford multiple simulations of multiple scenarios each over multiple decades. I share GE's concern that the existing model is already too cumbersome to be used effectively to study corrective measures. Indeed, this is probably the biggest model issue: as Yogi Berra might say, "If you can't use the model, you can't use the model." The following are some thoughts on the related issues of grid resolution and computational cost.

Response 4-EA-1:

The computational requirements of EFDC are significant, but EPA disagrees with the Reviewer's concern that the model is too cumbersome to be used effectively to evaluate remedial alternatives. See the General Response above.

Computational Costs

A number of options have been mentioned by the Peer Review Panel and by GE to improve model efficiency. These included splitting the calculation of hydrodynamics, sediment transport and PCB fate; running synthetic hydrological sequences; and employing alternate grid schemes. EPA did implement some efficiency measures (dynamic and split time steps, by-passing sections of the model that change slowly and running simulations of selected portions of the calibration period), but seemed to have dismissed the bigger ticket items. I think some of these have to be revisited.

There is no reason, in principle, why the model calculation of hydrodynamics, sediment transport and PCB transport/fate could not be de-coupled. In particular, the hydrodynamics could be run first. These results could be stored and used to transport sediment and PCBs. Since the hydrodynamics would only need to be computed once (hydrology does not depend significantly on scenario), considerable saving could be obtained. The devil is in the details of course, and this would probably take a few months of time. But since EPA is a model developer, and such a decoupled model would have future applications, EPA may want to invest in such a project, possibly calling in the model developer (J. Hamrick).

Response 4-EA-2:

The modeling team clearly recognizes the significance of run time issues associated with decadal-scale EFDC simulations, and acknowledges that the

> Reviewer has accurately summarized options mentioned by the Reviewers to improve run time. It is noted, however, had the various options summarized by the Reviewer been adopted for the model calibration, it would have been difficult to quantify the degree to which the approximations contributed to the residuals (measures of bias) between the model results and data. See the General Response above.
>
> In addition, while the development of the EFDC code for the Housatonic River application has resulted in benefits to further applications of the code (including within the Agency), generic model development is not a primary concern in this modeling study.

Further savings could be obtained using ***synthetic hydrologic flows.*** Based on the relatively large data base, a suitable discrete distribution could be generated that includes representative flows of different magnitudes and recurrence intervals. The model (with or without the hydrodynamics disaggregated from transport) could then be run for a synthetic year (or a short period of years large enough to include the largest/least frequent flow of interest), and the changes per year documented. Perhaps, without rerunning the model, ***the long term effects could be computed by extrapolation of the one year results.***

Response 4-EA-3:

> The application of the model in the CMS will be discussed in the Model Working Group and subsequently proposed by GE and reviewed by EPA. Screening techniques, such as that described by the Reviewer, may be considered.

Also, it still seems to me that in channel and over bank calculations could be de-coupled, saving additional computational time, and/or allowing more detail during the vast majority of time when the flow is within the banks and essentially one-dimensional. EPA claims there are coupling issues such as loss of momentum conservation, which may be true. But exact coupling is not critical. We are only expecting the model to predict gross trends in PCB concentrations (over time and the longitudinal direction), so a temporary mis-match might be acceptable. In general, I would support more a somewhat more approximate model (e.g., with parameterizations) that could be more efficiently used.

Response 4-EA-4:

> It has been estimated that 60% of the PCBs present in the PSA are in the floodplain, and risks to human health and the environment were identified from exposure to PCBs in the floodplain. This was an important factor in the decision to include the floodplain in the computational grid. Logic tests in EFDC provide a means of bypassing computations for floodplain cells that are not inundated, so the floodplain and channel are effectively decoupled during the majority of the time when in-bank flows are simulated. The Reviewer's suggestion of adding more detail to the in-channel grid for within-bank flow conditions would increase the computational burden for periods that are less important for solids and PCB transport. Evaluation of options, such as different grids for high-flow and low-flow

> conditions, or decoupling the hydrodynamic, sediment transport, and PCB fate submodels of EFDC, were evaluated. In these evaluations, an important consideration was whether the potential reduction in run time would shorten the calibration and validation tasks by more than the time required to develop and test the linkages or code modifications. See the General Response above.

In this regard, I like M. Garcia's idea of a flood plain number to quantify deposition of PCBs within portions of the floodplain. PCBs appear to be on a conveyor belt whereby they are resuspended from within the channel at high flow, partly deposited on the banks, then gradually eroded from the bank back into the river, until some of them are deposited in Woods Pond. These processes are all in the model, but it might be nice to document the life history of a "numerically marked" cohort of sediment, and to test the sensitivity of the ultimate fate of this cohort to processes such as the bank erosion rate, frequency of over bank flows, etc. By decoupling the sediment transport and fate, this could lead to a model simplification, whereby the over bank processes (deposition and bank erosion) are parameterized as simple first order sinks and sources whose rates are determined by calibration to detailed simulation over the calibration period. These rates could vary with flow conditions, but would be independent of PCB concentration, allowing simple calculations during the Corrective Measures Study.

Response 4-EA-5:

> A discussion of the limitations of Dr. Garcia's recommendation regarding a dimensionless floodplain number is provided in Appendix B.7 to the FMD and in the response to General Topic 3C. It is noted that the Reviewer's conceptual description of PCB transport does not mention PCB deposition onto the floodplain in locations distant from the river channel in locations where remobilization due to bank erosion is not possible. The Reviewer's concept of numerically marking a "cohort of sediment" has been simulated in the source–specific (tracer) simulations. For those simulations, the "numerically marked" sediment included all of the sediment in an individual reach, or in floodplain soil. Changes in input and/or output processing would permit a similar analysis at finer or coarser spatial scales.

In summary, it is a shame that the longitudinal and transverse grid sizes (Δx and Δy) must be similar since the former need not be nearly so small (fish average their exposure over large distances) and the latter should be smaller. In my previous comments I suggested eliminating Δx as a dependent variable (substituting, instead, grain size by reach), but I recognize this was a radical idea that might take too much time to implement.

Response 4-EA-6:

> The physical character of the Housatonic River presents a challenging setting for grid development and contaminant fate and transport modeling, which is further complicated by the need to include the floodplain in the computational domain. As a result of the meandering channel, the upstream-downstream direction is along the north-south axis in some places and along the east-west axis in others. In terms of the grid, this means that the main channel runs along columns of the

grid in some locations and along rows of the grid in other locations. This is the basis for the Reviewer's observation that "the longitudinal and transverse grid sizes (Δx and Δy) must be similar." The dimension that is transverse in one portion of the grid is longitudinal in another. It was not clear how the Reviewer's suggestion of eliminating Δx could be accomplished in light of the comments from other Reviewers that an even finer spatial representation of bathymetry was desirable to compute finer spatial resolution of bed shear stress. Comments regarding grid scale are addressed in the responses to General Topic 3.

Finally, while the above discussion refers to lateral resolution (and its relationship to computational cost), I am also concerned that the vertical resolution within the sediment bed that may be insufficient to resolve near bed gradients that might result from remedial efforts (see Question 1). Clearly, calculations in the sediment bed can use a different (much longer) time step than those in the water column. It is not clear how costly additional vertical resolution in the sediment bed would be, but I think it should be considered.

Response 4-EA-7:

The vertical resolution of the bed was based on an extensive effort that included the use of site-specific benthic taxonomic data, and a considerable literature review on the topics of mixing depth and mixing rates. In response to additional concerns about the representation of the bed, which were expressed by the Reviewers at the Document Overview Meeting and at the Peer Review, sediment profile imaging was conducted. These images support the model's vertical representation of the bed. Additional details on these topics are provided in FMD Appendices B.4 and B.5. Further responses are provided in General Topic 3D.

The Reviewer is correct in noting that calculations within the sediment bed can use a different time step than the water column. This option was invoked, and sediment calculations were performed at a multiple (16) of the time step used in the water column. Additional resolution in the vertical dimension in the bed was evaluated and substantial increases in run time were encountered because the increase in memory requirements forced the use of virtual memory.

W. Frank Bohlen

Even with these process questions resolved there remains the issue of run-time. It seems clear that the model as presently configured requires entirely too much time for the completion of a single run to be useful within timely evaluations of a significant number of remedial schemes. We probably knew this several years ago and should have been more sensitive to the need to develop alternative formulations. A number of these, including the separation of the hydrodynamics from the transport estimates and subsequent FCM evaluations were previously mentioned. It is now necessary and possible to go further. Using the experience gained from "whole model" runs it should now be possible to develop a number of synthetic hydrographs detailing streamflow, stage, and TSS concentrations at the upstream boundary and each of the primary tributary streams. This would eliminate the need to run HSPF on a regular basis. EFDC is a ponderous model and can be streamlined now that we have a better idea of the relative importance of the governing variables. John Hamrick should be charged with this task (no more than 6 months) as soon as possible. As part of this streamlining the model grid characteristics should be carefully reviewed, again using what has been learned about the relative importance of each of the domain regions (sidebanks, backwaters, floodplains etc.). My sense of the present grid is that it underspecifies the channel region and overspecifies the floodplain. The latter could almost be treated as a box with fluxes simply proportional to stage which could be specified along its margin by EFDC. If in time more detail of the interior of the plain is needed for remedial purposes consideration might be given to replacing the high resolution grid on the plain while placing a box in some other area (channel sidebanks?). Finally, I'd consider eliminating the FCM in favor of a parametric (flow, TSS concentrations?) approximation of body burden uptake based on the results of the complete model runs.

Response 4-FB-1:

EPA disagrees with the Reviewer's comment that the run time prevents it from being "useful within timely evaluations of a significant number of remedial schemes." As discussed in the General Response above, EPA believes that a considerable number of simulations can be executed concurrently, making the model a useful tool for the CMS. EPA disagrees with the Reviewer's suggestions to minimize the roles of the watershed and bioaccumulation models (HSPF and FCM, respectively), since these models have significantly smaller run times than EFDC and little time savings would be gained while considerable rigor would be lost. The Reviewer's suggestion to trade existing grid-resolution in the floodplain for additional resolution in the river channel ignores the potential impact that such a change would have on run times, which would be expected to result in an increase in run time. These factors are discussed in FMD Appendix B.1, and in Response to General Topic 3.

In short, what I'm thinking about is the development of an supplementary modeling scheme for use in the CMS. The complete model would serve as a guide assisting in the development of a series of simpler, more efficient but less comprehensive, formulations that would be directed at particular remedial schemes. The complete model framework would remain in place providing guidance regarding the need for and type of data to supplement model formulations for both calibration and verification purposes but would not be run as frequently as the supplementary schema. The alternative might be to turn to a different series of models entirely. This is not recommended without good reasons that I don't have at the moment.

Response 4-FB-2:

GE and their consultants will have the option of using techniques suggested by the Reviewer, or other methods to screen scenarios before running the full model. Such options will be discussed in the Model Working Group in the formulation of the CMS.

Douglas Endicott

The low-flow simulations have been criticized by GE as underpredicting the increase in PCB concentrations observed between the upstream boundary condition and New Lenox Road, as well as the decrease in PCB concentrations across Woods Pond. In addition, the profiles show a consistent "drop" in PCB concentrations immediately below the upper boundary. These features have been interpreted as suggesting that processes related to PCB fluxes between the water and sediment may be misrepresented or miscalibrated.

Response 4-DE-1:

The Reviewer's comments refer to different locations within the PSA. The consistent "drop" in water column PCB concentrations immediately downstream of the upper boundary reflects the effect of the West Branch inflow, which typically dilutes the PCB concentration from the East Branch. The following responses begin with comments related to the most-upstream location, and then progressively move to comments related to downstream locations.

EPA acknowledges that upstream boundary concentration estimates for PCBs affect the upstream portion of the spatial profiles of water column PCB concentrations; however, evaluations were conducted that indicate that this is not a concern (see General Topics 1 and 6, and Appendix B.2 to the FMD).

The issue of variability in water column concentrations on the ability to assess spatial patterns in PCB concentration across Woods Pond was discussed at the Document Overview Meeting. In response to this Reviewer's comment regarding spatial patterns across Woods Pond, a t-test was performed to assess the difference in the average concentrations entering and leaving the Pond, which are presented on the low-flow spatial profiles (MVR Figures 4.2-69 and 6.2-45). The confidence interval of the difference in mean concentrations includes zero, so at a 95% confidence level, the low-flow data do not indicate that the mean concentrations entering and leaving the pond are different. The concentrations simulated by the model also show little change across Woods Pond.

Finally, the increase in surficial sediment PCB concentrations over time within the upper-most sediment reach (RM 135.13-134.89) seems problematic and should be corrected. Is this behavior related to sudden drop in water column PCB concentrations downstream of the boundary condition, for example in the way bed load is initiated? Without access to the model, it is impossible to sort this out.

Response 4-DE-2:

In response to this Reviewer's comment on the artifact of the representation of bedload boundary conditions on sediment PCB concentrations in grid cells immediately below the confluence, a change was made in the f_{OC} of non-cohesive solids included in the upstream boundary condition to account for bedload solids. This change has eliminated the artifact noted by the Reviewer.

Marcelo H. Garcia

With the current computational mesh configuration, the application of the model is computationally taxing, resulting in extremely long execution times. This has made the calibration and validation of the model a very difficult exercise and will most likely affect the analysis of remediation alternatives in the future.

Response 4-MG-1:

While the Reviewer is correct in recognizing the impact of run times on the calibration/validation effort, recent computer hardware advances have resulted in the ability to execute runs in 50% to 67% of the MVR run times. These recent advances will provide significant benefit for the use of the model in the CMS.

While uncertainty in model inputs and outputs was recognized, the nature of the model does not allow for a conventional uncertainty analysis. Once all the processes are accounted for and well represented in the model and assuming that the EFDC code can be run more efficiently, it might be feasible to conduct an uncertainty analysis.

Response 4-MG-2:

The Reviewer recognizes a practical limitation that computationally intensive models impose on the feasibility of performing uncertainty analyses. Based on Comments from the Review Panel and discussions of the Model Working Group, EPA and GE have agreed to modify the approach for uncertainty analyses. This topic is discussed in more detail in responses to General Topic 5.

Frank Gobas

The spatial resolution of the environmental fate and food-web model is unnecessarily complex. A much simpler spatial model is consistent with the empirical data and the current model calculations. The hydrodynamic model does not run on a low spatial resolution. However, it is possible to run the hydrodynamic model on its optimal spatial resolution and aggregate the hydrodynamic data for a much simpler PCB environmental fate model. The latter would also reduce the model run-time from the current, unacceptable 30 to 50 d. This would also make the model more transparent due to greater simplicity. The modelers could argue that in terms of predicting spatially varying concentrations the current multi-compartment model is consistent with a much simpler lower spatial resolution model, so there is no need to develop the simpler model. This is true for the model’s current application. However, the modelers should keep in mind that when the current model is applied to make predictions of the impact of remedial actions on PCB concentrations as a function of space (i.e. model objective #1), the model may predict spatial differences in concentrations that have not been “validated” or “ground truthed”.

The modelers should consider developing a simpler spatial representation of the environmental fate and food-chain model by decoupling the hydrodynamic and sediment transport model from the environmental fate and food-chain model. The current high spatial resolution of the model is unnecessary for the environmental fate and food-chain model. Also, the time-step requirement for the hydrodynamic model is too onerous and unnecessary for the environmental fate and food-chain models. This change in model design will make it possible to make many model runs within a reasonable computational time.

Response 4-FG-1:

The Reviewer suggests that the results of finer-resolution grids for hydrodynamics (and presumably sediment transport) could be aggregated onto a coarser grid for PCB fate modeling. Implementation of this approach would require that, in addition to decoupling the hydrodynamics and sediment transport sub-models, development of a linkage scheme would be necessary to map erosion and deposition fluxes from the finer-resolution hydrodynamic and sediment transport grids onto the coarser PCB fate grid. These modifications could be accomplished and tested given time and resources. However, it is highly unlikely that the reduction in run time for the PCB fate component would offset the schedule delay introduced by the code modification effort, and the resulting model performance is unknown.

The run times quoted by the Reviewer are consistent with 60- to 100-year simulations, run on the computer hardware used during the Phase 2 calibration and validation modeling. Recent upgrades in computing technology have reduced the run times for these simulations to between 2 to 4 weeks.

When the model is applied to the evaluation of remedial alternatives, spatial patterns of predicted PCB concentrations may very well represent conditions that could not be ground-truthed because they haven’t existed in the river due to the ongoing input of PCBs from the East Branch. This is often the case when models are used to simulate a response to a change in forcing that has not occurred, and, in fact, this is the reason that such models are developed. The model calibration and validation process was implemented to test the parameterization of the mathematical representations of processes that relate model forcing and responses over space and time, as documented by data collected in the past 26 years. The calibration strategy (see FMD Section 3.3 and the response to General Topic 1) adopted for this study relied on site-specific data, when available, for estimation of model parameters, which were tested by focusing on conditions when a limited number, or only one process, dominated the transport of solids or PCBs. By calibrating individual processes to the extent possible, confidence is derived in the model’s ability to represent the relevant processes under a different set of forcing, such as lower sediment PCB concentrations.

Reduce the model run-time from an unacceptable 30 to 50 days to 1 d (at most), such that different model parameterization schemes can be explored for making model projections.

> **Response 4-FG-2:**
>
> While a model run time of 1 day is highly desirable, and EPA has explored many different recommendations from the Peer Reviewer Panel or the Model Working Group, none of these approaches will realize the goal of a 1-day run time.

Wilbert Lick

Numerical Gridding

With the present grid, the width of the river is generally approximated as one cell. In the Housatonic, as in most rivers, there are large differences in erosion between the deeper and the shallower parts of the river. Predicting the dissimilar amounts of erosion/deposition across a cross-section of the river is crucial in predicting the long-term exposure of PCBs by erosion and/or natural recovery by deposition. Averaging across the cross-section does not describe the erosion/deposition process accurately. A minimum of three cells across the river (two shallow, near-shore cells and one deeper, center cell) should be used.

In the floodplain, our knowledge of the basic processes of erosion/deposition and the non-erosional/depositional flux is poor. Because of this, a very coarse grid can be used to approximate the processes in this area.

A better description of the bathymetry of the river will increase the computational time. Drastically decreasing the number of grid cells in the floodplain will significantly decrease the computational time. The computational time can also be decreased by (a) separating the hydrodynamics, sediment transport, and contaminant transport calculations and (b) for small and moderate flows, approximate and calculate the hydrodynamics and transport as a sequence of steady-state solutions and only treat big events in detail.

> **Response 4-WL-1:**
>
> The Reviewer’s generalization about large differences in erosion in different parts of a cross-section is not uniformly consistent with data obtained through repeated surveys of 9 locations in Reaches 5A and 5B (see FMD Appendix B6). The Reviewer’s comments regarding the lateral grid resolution are addressed in General Topic 3. The Reviewer’s assumption that the increase in run time associated with increasing the grid resolution of the channel can be offset by drastically reducing the number of grid cells in the floodplain is incorrect. Logic tests in EFDC provide a means of bypassing computations for floodplain cells that are not inundated, so the floodplain and channel are effectively decoupled during the majority of the time when in-bank flows are simulated. Increasing the number of grid cells in the channel would increase the time required for computations in channel cells under all flow conditions.

The calculations of the hydrodynamics, sediment transport, and PCB transport should be separated.

Response 4-WL-2:

The modeling team considered options, such as decoupling the hydrodynamic, sediment transport, and PCB fate portions of EFDC. An important factor in the evaluation was whether the potential reduction in run time would have shortened the calibration and validation tasks and CMS study by more than the time required to develop and test the linkages or code modifications. The approach described in the MVR and FMD was followed because it was clear that the proposed options would not provide a benefit that would outweigh the potential cost in terms of dollars and delay in the schedule.

5. GENERAL TOPIC 5: SENSITIVITY AND UNCERTAINTY

5.1 GENERAL TOPIC 5A: SENSITIVITY

5.1.1 Comment Summary

Reviewer Name	Page Number/Line Number
Adams	Page 9, lines 18 to 31
Bohlen	Page 12, line 28 to page 13, line 3
	Page 13, lines 5 to 11
Endicott	Page 16, line 19 to page 17, line 18
Garcia	Page 5, line 39 to page 6, line 2
Gobas	Page 9, lines 1 to 21
	Page 14, lines 15 to 19
Lick	Page 1, lines 37 to 39
List	Page 5, lines 31 to 34

5.1.2 General Topic Summary

The sensitivity of model output to a particular input parameter is defined as the ratio (expressed as a percentage) of the average absolute percent change in model output for the two model runs, to the average absolute percent change in input parameters. Selected model input parameters were varied individually and the response of selected model output variables was noted and compared to the base case (e.g., the final Phase 2 Calibration). Additional types of analyses were also used to investigate model sensitivity to factors that did not lend themselves to the formal sensitivity approach.

Reviewers had differing opinions on the sensitivity analysis presented in the Model Validation Report. Four of the Reviewers commented that the sensitivity analysis was adequate or good, although most provided suggestions for additional analyses or evaluations that could or should be conducted to improve the sensitivity analysis. Three Reviewers believed that the sensitivity analysis was inadequate because the basic model processes had not been adequately characterized and such characterization was a necessary first step. One Reviewer added that it was not possible to conduct a proper sensitivity analysis due to the coarseness of the model grid.

Two Reviewers specifically questioned an apparently anomalous result that indicated limited sensitivity of PCB concentrations in Woods Pond to parameters controlling diffusion of contaminants from the sediment bed.

5.1.3 General Topic Response

EPA agrees with the majority of the Reviewers, who commented that the sensitivity analysis presented in the Model Validation Report provides a reasonable characterization of model sensitivity. EPA appreciates the Reviewers' suggestions regarding additional sensitivity analyses that could be informative.

Although not conducted specifically as part of the sensitivity analysis, the effect of various combinations of "shutting off" PCB sources was thoroughly investigated in the process-based flux summaries (tracer runs) discussed in the FMD. Another additional type of sensitivity analysis that was performed outside of the formal approach presented in the MVR was to evaluate the effect of different grid cell sizes during model development, in the MFD, and in Appendix B.1 to the FMD.

EPA believes that the sensitivity analyses conducted provide sufficient assurance that the model(s) as currently structured and parameterized are responding correctly to the forcing functions in the system. Although other analyses might assist in further understanding model behavior, they are not necessary to achieve the purpose of the modeling study.

With regard to the comments concerning the inability to conduct sensitivity analysis due to errors in characterization of model processes or grid resolution, such comments are not comments on the sensitivity analysis per se, but rather on other aspects of the modeling study that are discussed elsewhere in this Responsiveness Summary and in the FMD. The sensitivity analysis presented in the MVR is an evaluation of the sensitivity of the model framework as currently structured and parameterized, and comments regarding the validity of that structure and parameterization are addressed under other General Topics and are not further addressed here.

5.1.4 Response to Specific Comments on General Topic 5A

Eric Adams

Most sensitivities seem reasonable. As expected, flows show strong sensitivity to upstream boundary flows, TSS is sensitive to upstream boundary flow and TSS, and PCBs are sensitive to upstream boundary flow, TSS and PCB concentration.

At low flow PCB there is strong sensitivity to K_f and partitioning (at least at New Lenox Rd), which makes sense, since "diffusion" is the only important process, but this effect is largely diluted out in the 10.5 year simulation. At high flow, there is strong sensitivity to parameters dealing with particulate phase transport (settling, erosion).

While the model shows substantial sensitivity to diffusion parameters during low flow at New Lenox Rd, GE points out that virtually no sensitivity is indicated at Woods Pond Footbridge. Since the flow must pass from NLR to WPF, this is illogical. Unless there is a mistake in plotting (and EPA suggested there was not), this could indicate a potential problem in model formulation.

Response 5A-EA-1:

EPA agrees with the Reviewer that most sensitivity results seem reasonable. The exception noted by the Reviewer and by GE for the low-flow tornado diagrams at Woods Pond Footbridge was caused in the post-processing that extracted the model results; this error has been corrected. The revised tornado diagram for PCB fluxes at Woods Pond Footbridge, which is included in Section 3.7 of the FMD, shows sensitivities for the low-flow period that are similar to the results for New Lenox Road.

W. Frank Bohlen

In most cases the sensitivity of each of the models to the significant state and process variables has been adequately characterized. However, as discussed above, I remain concerned about the accuracy of the sediment transport processes within the model domain and look to the sensitivity analysis for some guidance. The sensitivity analyses provided in the MVR and the earlier calibration report provide clear indication that sediment mass flux throughout the PSA is very sensitive to the upstream boundary conditions. With these reasonably well specified, using a combination of HSPF results and field observations, it may not be very difficult to achieve reasonable model/data agreement despite inaccurate simulation of bed erosion within the PSA. The sensitivity analyses do not take the next step to test the adequacy of the interior formulations by shutting off the boundary contribution of sediment and PCB's. The importance of this formulation will progressively increase as the upstream sources of contamination are reduced and remedial measures for the downstream areas are being evaluated. A detailed analysis of the sensitivity and accuracy of this formulation is recommended The results of this effort might serve to address the concerns raised in GE's review of the MVR dealing with the relative importance of Reach 5A and 5B as sources of sediments and contaminants to Reach 5.

Response 5A-FB-1:

The Reviewer is correct that the sensitivity analyses indicate the sensitivity of EFDC to the solids and PCB loads entering the domain at the upstream boundary. However, EPA does not believe that the sensitivity analysis provides the best tool to determine if such loads have been adequately characterized.

The absolute and relative magnitude (boundary loads versus internally generated through erosion) of solids fluxes within the PSA can be evaluated more directly from the output of the example scenario and tracer simulations (MVR Figure 4.2-51, FMD Figures 4.2-1 and 4.2-2) rather than through sensitivity analyses in which upstream sources are set to zero. Because the settling velocity of cohesive solids is dependent on the cohesive solids concentration, sensitivity analyses with a zero upstream boundary condition will alter the settling velocity applied to the suspended solids, which in turn will affect the transport of solids through the system.

Model output from the Phase 2 calibration simulation was processed to summarize solids fluxes from upstream boundaries and sediment erosion between various locations in the river for a range of flow conditions (MVR Figure 4.2-51). The information presented in this figure can be used to compute the ratio of erosion fluxes to upstream boundary inputs, and how the ratio changes with river flow. The process-based flux summaries for Phase 2 calibration and validation simulations indicate the magnitude of solids (FMD Figures 4.2-1 and 4.2-2) and PCB fluxes (FMD Figures 4.2-6 and 4.2-7) from various sources and processes. The spatial pattern of erosion fluxes and the variation in the flux rate with flow are consistent with the conceptual model in that more erosion is calculated in the higher energy upstream areas and erosion increases dramatically with flow.

In addition to testing of the erosion formulation additional sensitivity analyses should be performed to assess model response to the thickness of the active sediment column. This has been a matter of concern for some time. Experience in other riverine/bayou systems as well as the detailed core data (Fig. 4-21c, attached) indicate to me that the active bed used in the model is too thick. It may be that this specification in terms of physical transport characteristics has relatively little effect on overall model results (although that might bring up another set of questions). A test of model response to this characterization is recommended.

Response 5A-FB-2:

The thickness of the surface sediment layer, as well as the general structure and dynamics of the entire sediment bed, were extensively evaluated at different stages throughout the model development process and a number of different configurations were tested. Based on the results of that testing and evaluation, the sensitivity of the model to variations in the thickness of the sediment surface layer is well understood by the EPA modeling team, and it was apparent that the model behavior was reasonable after performing these tests and also in evaluating the balance of fate processes. The issue of surface sediment layer thickness is not one of sensitivity per se, but rather of whether the thickness of the surface layer in the model properly represents the actual configuration in the river (see Appendix B.5 to the FMD).

Douglas Endicott

The MVR does a good job of illustrating the sensitivity of the EFDC model predictions to input parameters. Several additional model inputs should be added to the sensitivity analysis, and these are identified below. As pointed out by GE, there are also some sensitivity analysis results that are counter-intuitive; these should be explored and explained.

Sediment transport sensitivity analysis is displayed in terms of peak and mean TSS fluxes, for the entire calibration period as well as 2 different flow events (Figures 5.1-3 thru 5.1-5). It would be useful to also see the sensitivity of the model displayed as changes in bed elevation.

Part of the sensitivity analysis for the sediment transport model should include the sensitivity to the grid resolution because, as discussed previously, the prediction of sediment bed scour may depend on the number of lateral cells in the channel cross-section. This is only needed for a “test reach” of the river channel, preferably the reach where bathymetry transects were measured. Note that Earl Hayter’s previous analysis was not informative on this issue.

PCB fate and transport sensitivity analysis is displayed in terms of peak and mean PCB fluxes, again for the entire calibration period as well as the two flow events (Figures 5.1-6 thru 5.1-11). The sensitivity of the Woods Pond PCB flux predictions at low flow (Figure 5.1-9) is puzzling, since it suggests practically no sensitivity to parameters associated with sediment-water exchange of PCBs. While there may be little exchange within Woods Pond itself, considerable PCB exchange in upstream reaches should be reflected in the sensitivities of these parameters in Woods Pond. Instead, Figure 5.1-9 show that low-flow PCB fluxes in Woods Pond are only mildly sensitive to the upstream boundary inflow, and really nothing else. It remains for EPA to explain how this result is evidence that “the model is properly representing the physics of the river as it relates to the movement of solids and contaminants”.

As stated previously, the initial PCB concentrations specified for the sediment bed should be added as parameters in the sensitivity analysis, for all reaches where they are not independently determined from data. Likewise, the surficial sediment layer thickness should be included as a parameter in the sensitivity analysis.

Response 5A-DE-1:

EPA agrees with the Reviewer’s comment that the results presented in the MVR illustrate the sensitivities in the EFDC parameterization. Between the Model Calibration Report and the Model Validation Report, the sensitivity simulations were increased to the 10½-year Phase 2 Calibration period, and additional parameters were included in the analysis. A decision was also made to show the sensitivity of the response of each EFDC submodel to parameters tested for all preceding submodels (e.g., sensitivity of PCB fluxes to the hydrodynamic parameter for bottom friction, z_0). In addition, an alternate presentation format for the results, referred to as a tornado diagram (see Section 3.7 of the FMD), was adopted to allow a consistent presentation of the results of the sensitivity analysis for all three models. In the interest of focusing the EFDC sensitivity analysis on the output variables of most importance, the food chain exposure concentrations were added as response variables because they synthesize the effect of several EFDC variables, and represent the most important output of the EFDC simulation.

The effect of grid resolution was investigated as part of the MFD phase of the modeling study. The suggestion of performing sensitivity analyses for the grid resolution was not implemented because this analysis was estimated to involve a substantial effort that could provide only limited benefit beyond the analyses that had been previously performed. Simply applying all of the calibration parameters on a more refined grid could lead to inappropriate and uninterpretable comparisons. For example, sub-grid-scale features in the current grid, such as undulations in the bed elevation, have been subsumed into the parameterization of the effective roughness height (z_0). If the resolution of the grid was increased to explicitly represent these features, a reassessment of the appropriate z_0 would be required to produce the same hydrodynamic transport through the new grid. Data would be required to develop boundary conditions for a laterally segmented grid, and the length of a “test reach” would have to span approximately 5 miles, to include locations where monitoring data, including water-surface elevations, were collected. Obtaining different results from a refined grid of a “test reach,” compared to the results for the same reach from the original grid, would not be meaningful because there would be no basis for concluding that one result was better than the other.

A valuable check on the calculations for the current grid applied in Reaches 5 and 6 is the simulated sedimentation rate in Woods Pond, where model-data comparisons of sedimentation rates are available for 12 locations. This check would not be available for simulations performed in only a portion of the domain. Because EPA and GE concur that lateral segmentation of the river channel is not necessary and that the current grid is appropriate for use in the CMS, the effort required to perform the recommended sensitivity analysis is not justified.

The Reviewer recognizes that data were too limited to develop sediment PCB initial conditions for the Phase 2 calibration period for the the validation period (except for Reach 6). EPA recognizes that the use of the model to make hindcast estimates of initial conditions for times when data are not available does not eliminate the uncertainty resulting from lack of data. In the sensitivity analysis, model inputs were varied by ±50%. Based on the temporal trends in sediment data and fish tissue concentrations (see Appendix A.1 of the FMD), a variation of ±50% in sediment initial conditions would greatly overstate the likely range in sediment concentrations and could be misinterpreted as an indication of uncertainty.

In response to comments from the Peer Reviewers, sediment profile imaging (SPI) was conducted in early September 2006 to obtain additional site-specific information on the structure of the sediment bed of the Housatonic River. The information obtained in the SPI work supported the bed layering used in the Phase 2 Calibration and Validation modeling (see FMD Appendix B.5 for additional information). Based on the conclusions drawn from the site-specific information, and considerable investigation of the response of the model to surface sediment layer thickness during model development, the thickness of the sediment layers was not added as a parameter in the sensitivity analysis.

EPA's assessment that EFDC properly represents the transport of solids and PCBs (including in Woods Pond) was based on the evaluation of more extensive model output than the graphical presentations of the sensitivity analysis results. The post-processing that extracted the model results for the low-flow tornado diagrams contained an error, which has been corrected. The revised tornado diagram for PCB fluxes at Woods Pond Footbridge shows sensitivities for the low-flow period that are similar to the results for New Lenox Road.

Marcelo H. Garcia

Given the facts that several processes are yet to be fully characterized in the model and that a coarse computational grid has been used for the model calibration, I do not think that a full sensitivity analysis can be conducted at this time.

Response 5A-MG-1:

EPA disagrees that questions raised by Reviewers regarding the suitability of model structure and/or process formulations should preclude assessment of the model sensitivity. As noted above, the sensitivity analysis presented in the MVR is an evaluation of the sensitivity of the model as currently structured and parameterized; questions regarding the validity of that structure and parameterization, including particularly the appropriateness of the computation grid used for the modeling study, are addressed under other topics in this Responsiveness Summary and in the FMD.

Frank Gobas

The modelers have done a good job describing parameter sensitivities. The latter is not easy as there are many parameters and the sensitivities of each of them depend on the values chosen for the others.

However, there is one area in the model where conducting a thorough sensitivity analysis is crucial because of the lack of good calibration data and difficulties in model parameterization. This is in the component of the Environmental fate model that controls the long term temporal response of the model. The selection of the thickness of sediment layers, diffusion rates, and bioturbation rates are all very difficult. Hence, a thorough sensitivity analysis is crucial to determine the bounds within which temporal changes in concentrations can be expected to occur. This analysis is not presented in the current reports and supports my conclusion that the long term temporal response of the model is too uncertain to make meaningful predictions.

Recommendations:

Given the lack of calibration data, the sensitivity analysis is probably one of the few things that can be done to increase confidence in any estimates of the temporal response in PCB concentrations calculated by the model. I recommend that it is considered and added. For this sensitivity analysis to be doable and convincing, the modelers could just focus on the part of the model that controls the long term temporal response of the model.

Response 5A-FG-1:

In response to comments from the Reviewers, additional parameters were added to the sensitivity analyses, so that the sensitivity analyses include all but one of the parameters listed by the Reviewer (see FMD Section 3.7). The one exception is related to the thickness of the surface layer of the sediment, and is discussed above in Response 5A-FB-2.

In response to comments from the Reviewers, sediment profile imaging (see Appendix A.3 to the FMD) was conducted in early September 2006 to obtain site-specific information on the structure of the surface sediment of the Housatonic River. The site-specific information obtained from that effort supported the bed layering used in the Phase 2 calibration and validation modeling (see FMD, Appendix B.5 for additional information).

Conduct sensitivity analyses with the goal to (i) further investigate the parameterization of the sediment-water exchange of PCBs, on which the temporal response of PCB concentrations in the River largely depends and (ii) improve the parameterization of the key processes if possible (see under charge question #1 for additional details on this issue).

Response 5A-FG-2:

In response to comments from Reviewers, additional parameters were added to the sensitivity analyses (see Section 3.7 of the FMD). Previously, the sensitivity analyses included the sediment-water mass transfer coefficient, which is applied to pore water PCBs, and the following parameters that affect the sediment-water column transport of sorbed PCBs: cohesive solids settling velocity and critical shear stress for deposition of cohesive solids; non-cohesive solids grain size, which affects both the settling velocity and resuspension of non-cohesive solids erosion; the critical shear stress for cohesive solids erosion; and the cohesive solids erosion rate. Additional parameters that were added include the exponent for the sediment particle mixing rate, cohesive solids erosion formulation, and the non-cohesive solids erosion rate.

Wilbert Lick

The sensitivity of the model to the parameterization of the significant state and process variables has not been adequately characterized. See discussion in the complete response.

Response 5A-WL-1:

The Reviewer's comment does not address model sensitivity but instead repeats concerns discussed under General Topic 1, which relates to the balance of fate processes. EPA disagrees that differences of opinion regarding the suitability of model process formulations should preclude assessment of the model sensitivity, and notes that the sensitivity analyses as conducted relate specifically to the model as currently structured and parameterized. See also responses to General Topic 1.

E. John List

My opinion is that the underlying processes involved with the PCB fate and transport do not appear well enough understood for a meaningful model to be developed. If the processes that seem to define the PCB fate and transport are not included in the model then their sensitivity cannot be assessed.

Response 5A-JL-1:

See Response 5A-WL-1.

5.2 GENERAL TOPIC 5B: UNCERTAINTY

5.2.1 Comment Summary

Reviewer Name	Page Number/Line Number
Adams	Page 9, line 36 to page 11, line 16
Bohlen	Page 13, line 13 to page 14, line 2
Endicott	Page 1, lines 12 to 26
	Page 17, line 22 to page 18, line 22
Garcia	Page 6, lines 13 to 17
Gobas	Page 10, line 34 to page 11, line 41
	Page 13, lines 27 to 31
	Page 14, lines 1 to 9
Lick	Page 1, lines 41 to 42
List	Page 5, lines 39 to 40

5.2.2 General Topic Summary

Methods for the quantitative evaluation of modeling study results are in various stages of development and there is no commonly accepted way of quantifying or communicating to decisionmakers the accuracy of model predictions. Uncertainty is often discussed only qualitatively, or ignored, and beyond recognition of the fact that there is some measure of uncertainty in all predictions, uncertainty results are not used in making decisions based on model results.

Reviewers were sharply divided on the validity and utility of the uncertainty analysis presented in the MVR. One Reviewer described the approach to the analysis as “rather traditional,” but questioned the assumption of independence among input parameters, noting that it resulted in very large uncertainty bounds. Two other Reviewers, however, described the analysis as a “real attempt to develop new methodology” and “unprecedented,” and one of the Reviewers noted that the assumption of independence among parameters was common practice. One Reviewer described the uncertainty analysis as “seriously flawed.” Two Reviewers felt that because the model processes had been incorrectly described, no valid uncertainty analysis was possible, and two Reviewers felt the complexity of the model(s) did not allow for a typical probabilistic uncertainty analysis. There was some consensus among the Reviewers that, whatever the validity of the current uncertainty analysis, it was difficult to apply the results in a way that would inform management decisions made based on model output.

5.2.3 General Topic Response

Given that methods for the quantitative evaluation of modeling study results are in various stages of development and there is no commonly accepted way of quantifying or communicating to decisionmakers the accuracy of model predictions, EPA anticipated that the Reviewers would have a range of opinions regarding the validity and utility of the uncertainty analysis presented in the Model Validation Report, and in fact requested input from the Reviewers on an appropriate approach to evaluating model uncertainty during the Peer Review of the Model Calibration Report. Most modeling studies deal with uncertainty only qualitatively, and beyond recognition of the fact that there is some measure of uncertainty in all predictions, uncertainty results are typically not used in making decisions based on model results. The approach presented in the MVR was intended to provide a methodology that could provide useful information to those charged with making decisions based on model results. EPA does not suggest that this is the only useful way of evaluating and presenting uncertainty and did not intend for the approach to be prescriptive for future uses of the model.

In the MVR, EPA presented a probabilistic approach to evaluating and quantifying model uncertainty. The approach featured the ability to propagate uncertainty through the framework of linked models, e.g., the uncertainty in model results output from HSPF that were used as input to EFDC, and the uncertainty in model results output from EFDC that were then used as input to FCM, were maintained.

In subsequent discussions with the Model Working Group following the Peer Review of the MVR, it was agreed that the uncertainty of the model results in the evaluation of the relative performance of remedial alternatives would be addressed using the best professional judgment of the Model Working Group. Responses to the specific issues raised by Reviewers as summarized in this General Topic are provided below.

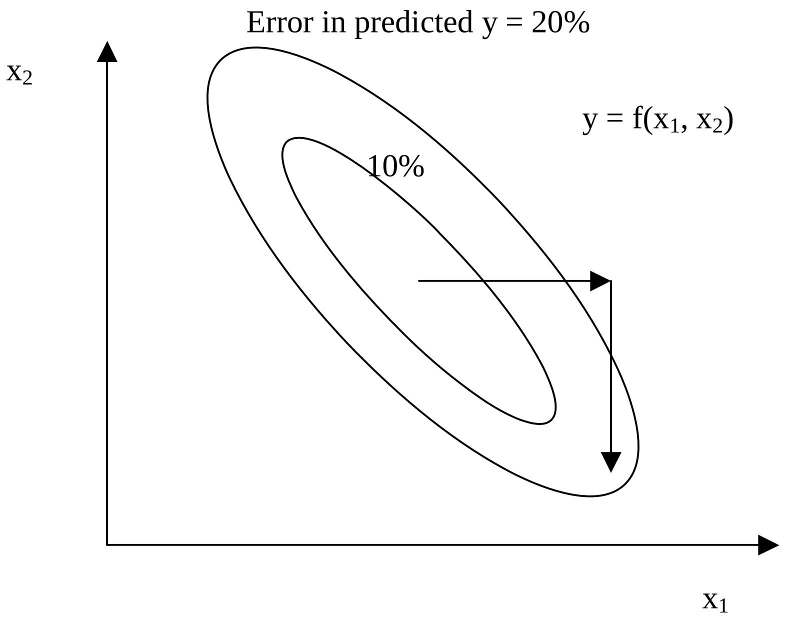
5.2.4 Response to Specific Comments on General Topic 5B

Eric Adams

In general, model uncertainty comes from two sources: imperfection in the basic model(s) (including unknown unknowns) and imperfection in the selection of the (known) model parameters. The former has been discussed to a certain degree under Question 1, which relates to processes, so the following relates mainly to the latter issue.

Although they use some sophisticated approaches, the modelers performed a rather traditional uncertainty analysis, varying known model parameters largely ***independently*** in a simulation of ***existing conditions***. If the model and data were both perfect, and the only uncertainty involved the values of model parameters, one could expect most of the parameters to be independent. (Some parameters that have a physical relationship, such as erosion parameters would still be correlated.) But the model has been calibrated to observed data, and the calibrated model parameters reflect both imperfection in the model (e.g., a parameter being used to cover more than one real process) and imperfection in the data (e.g., due to sampling or analytical error, insufficient sample resolution, etc). Thus, during uncertainty analysis, when one parameter is changed, other parameters should also change (i.e., they are not independent), as is illustrated in the following sketch. Here the dependent variable y is a function of two independent variables x_1 and x_2 and the contours are plots of some measure of error between measured and predicted values of y. If, as part of an uncertainty analysis, variable x_1 were changed (horizontal arrow), one would expect a correlated change in x_2 (vertical arrow) in order to minimize the error between measurements and predictions. Failure to take this correlation into account tends to overestimate the uncertainties in the model prediction. And, as suggested in the following paragraph, I suspect this issue becomes more extreme when the future conditions being predicted are substantially different from those under which the model was calibrated.

Apart from the question of parameter independence, there is the question of whether there are certain variables to which the model is more or less sensitive as a function of the remediation scenario. For example, following remediation there may be thin patches of clean/dirty sediment overlaying thicker regions of dirty/clean sediment, making calculations more sensitive to diffusion and partitioning than they were in the base case scenario. If this is the case, we may get a distorted impression of the effectiveness of remediation.



Based on sensitivity analysis, EPA tested the uncertainty of model outputs (e.g., PCB concentrations in water, bed and fish) to uncertainties in model input parameters. As GE points out, the model uncertainty is substantially greater than the uncertainty in the model (e.g., the s.e.m.) If the goal were simply to predict the uncertainty in present conditions, the uncertainty in the individual parameters could be scaled back until the uncertainties in the model and data matched. But it is also important to see the ***sensitivity of predicted model changes due to possible remedial alternatives*** to these same model input parameters. This is not difficult and should be performed either before or after the model is transferred to GE.

Finally, one use of the model is simply to compute relative fluxes, as EPA as demonstrated recently. Remedial decisions could be made on the basis of how much less PCB would be transported, rather than on (or in addition to) the basis of quantitative predictions of concentrations. Thus it would also be nice to see the ***sensitivity of predicted fluxes to the same model parameters***.

In summary, because of the lack of parameter correlation in the sensitivity study, and the fact that predicted uncertainty exceeds the uncertainty in the existing data, I believe the predicted uncertainty, when applied to future remedial conditions, will be too large.

Response 5B-EA-1:

It is true that correlations among input parameters can cause the output distributions from a Monte Carlo analysis to deviate from the true underlying distributions. However, it is not true that failure to quantitatively represent this correlation necessarily results in overestimation of the uncertainties in the model prediction. The effect of correlated inputs may result in either overestimation or underestimation of uncertainty, depending on whether the correlation is positive or negative, and also on the functional form of the relationship. Typically, uncertainty will be overestimated if positive correlations between variables that are added or multiplied are not included, or if negative correlations between variables that are involved in subtraction or division are not included; not including the reverse relationships will result in underestimation of uncertainty.

Specification of correlations among EFDC input parameters for the uncertainty analysis was considered; however, no reliable basis was available to quantify potential correlations, and, therefore, no intercorrelations were specified. Some of the process descriptions in EFDC naturally result in correlations between EFDC input parameters and simulated PCB concentrations (e.g., water column K_{OC} and water column sorbed PCB concentrations). These intercorrelations were included in the response surface model.

In the FCM Monte Carlo analysis, it was possible to specify the nature of some correlations, and when such intercorrelations could be reliably specified they were incorporated into the model. For example, for feeding preference assignments, when the dietary contribution of a given item was adjusted, the proportions of other dietary items were adjusted, such that the total amount ingested remained constant. This approach is consistent with the sensitivity and uncertainty assessment used in other published modeling applications. Burkhard (1998), in a sensitivity/uncertainty assessment for two bioaccumulation models, also applied a local sensitivity analysis method (generally assuming independence of inputs) but adjusted some parameters for known interdependencies. In that study, as with the FCM application, neither the lack of complete parameter independence nor the use of professional judgment in establishing input distributions was considered sufficiently problematic to render the results unusable.

EPA believes that comparison of the computed uncertainty bounds to the variability observed in the field data is an inappropriate means of determining whether the uncertainty bounds are plausible. There are four principal reasons for this opinion:

(1) The model, as a simplification of the complex natural processes governing transport and bioaccumulation, contains model uncertainties (including both uncertainties in represented processes/parameters and “unknown unknowns”) that would not be included in the field data.

(2) The field data contain natural variability over spatial and temporal scales finer than the intended scale of the model. For example, the fish tissue data include data on individual fish that have differing physical/biological properties and differing exposure patterns among individuals, whereas the model simulates the average condition within each reach.

(3) The field data incorporate analytical uncertainty (e.g., laboratory variability) that cannot be represented by the linked fate, transport, and bioaccumulation model.

(4) Field data are not available or are limited in number for some of the reaches and/or species of interest in the modeling study. Where insufficient data are available to assess model performance, model uncertainty evaluation cannot rely solely upon model versus data comparisons.

In summary, direct comparison of model uncertainty to data uncertainty is an inappropriate analysis, and cannot be used to assess the reasonableness of the computed uncertainty distributions from the linked Monte Carlo analysis.

W. Frank Bohlen

The uncertainty analyses presented in the MVR are interesting and represent a real attempt to develop new methodology for the assessment of complex models such as EFDC. Both Kolmogorov-Smirnov (KS) and Response Surface Modeling (RSM) analyses were applied. The KS methods require numerous model runs to develop a statistically robust data set. Given the time required for each run this was an onerous task and one made no easier by the failure of four of the EFDC runs. This effort yielded an initial estimate of uncertainties in the model prediction of PCB concentrations both water column and sediment associated. These concentrations were emphasized since they are passed to the food chain model (FCM). The KS analysis was supplemented by RSM to evaluate model uncertainty due to uncertainty in input parameters including flows, roughness parameters, critical erosion stresses, etc. The evaluations include consideration of the effects of the calculated uncertainties on the FCM predictions.

Overall the results of the uncertainty analyses, while interesting, were difficult to interpret in terms of model abilities to accurately simulate PCB transport and fate. The presentation of a series of individual summary figures for each condition does not facilitate interpretation in the absence of detailed description of cause and effect. The majority of the supporting text dealt primarily with methodolgy rather than interpretation. It may be that this type of analysis and its sophistication is premature and requires a greater understanding of the processes included in the model and their interactions than we presently have. As this becomes available with model use a better description and analysis of uncertainty might be possible.

Moving from the formal analyses of uncertainty to the general subject of model uncertainty, this report (MVR) too often fails to provide detailed discussion of uncertainty including consideration of causes. Statements such as “The simulated hydrographs....reproduced measured hydrographs reasonably well, however in some cases the magnitude of the simulated flow differed from the data in both positive and negative directions” (see pp 6-35) are too common. Uncertainty is to be expected as is variability both due to the input data being used and numerical model response. It should be introduced early in model discussions and included in logical discussion of cause and effect relationships throughout the report. I would hope that any summary report considers this as an absolute necessity. It will be of great value in building the confidence of a general readership.

Response 5B-FB-1:

EPA agrees with much of the Reviewer's assessment of the uncertainty analysis. The uncertainty approach presented in the MVR was intended to be demonstrative rather than prescriptive, and EPA recognizes that other approaches to evaluating model uncertainty may also be valid. In subsequent discussions with the Model Working Group following the Peer Review of the Model Validation Report, it was agreed that the uncertainty of the model results in the evaluation of the relative performance of remedial alternatives would be addressed using the best professional judgment of the Model Working Group. During the development of the FMD, further effort was made to provide discussion of the uncertainties associated with model outputs and data relationships (see Sections 3 and 4 of the FMD).

Douglas Endicott

I do have several comments to make regarding the peer review process itself. In the *EPA Response to Questions from Model Validation Document Overview Meeting*, EPA correctly notes that models are uncertain and that a degree of uncertainty is inherent in "these types of studies"(question: What scientific study does not include a degree of uncertainty?). They go on to identify three principal sources of model uncertainty (models are simplifications/approximations of reality, parameters are uncertain, future conditions are unknowable). Although these can be significant sources of uncertainty, EPA chose not to highlight several others, which may be equally important:

- Inappropriate process representations/formulations,
- Errors resulting from discretization of the model or the scale of its application,
- Errors in programming, preprocessing, running, linking and postprocessing the models, and
- Untested model predictions.

These are "lurking" sources of model uncertainty, which are not typically revealed by uncertainty analysis. I think it is useful to acknowledge these pitfalls, and keep them in mind throughout the course of a project. I think much of the Peer Review deliberations, and more than a few of the comments, reflect concern about these factors.

Response 5B-DE-1:

EPA agrees that the sources of uncertainty identified by the Reviewer are all potentially relevant to this (and any other) modeling study. These uncertainties are discussed individually below:

- Inappropriate process representations and formulations—This type of uncertainty was acknowledged by EPA in comments (dating back to the initial MFD) that models are a necessary simplification of reality. In conducting a phased approach to model development, from the MFD through Validation, and by engaging an independent Peer Review Panel at each major step of model development, EPA has consistently attempted to maximize the realism of the model processes and formulations. The model processes can never be “correct” because they are approximations of reality; however, they can be appropriate in the sense that they are useful, practical, and scientifically defensible representations of the environment that incorporate site-specific understanding of system behavior. To this end, the model evaluation is centered on determining “whether the simulation model is an acceptable representation of the real system, given the purpose of the simulation model” (Kleijnen, 1999). To this end, EPA believes that “error” in the specified processes stems primarily from necessary approximation of reality and not from erroneous application of processes. Further discussion of the balance of fate processes in EFDC is provided in the responses to General Topic 1 and in the FMD.

- Errors resulting from discretization of the model or the scale of its application—EPA acknowledges that errors can occur due to the need to discretize the model over space and time. EPA has minimized the effect of these potential errors by identifying relevant spatial and temporal domains and levels of detail commensurate with the needs of each model. The selection of a grid system, computational time step, and aggregation of data for model linkages have all been rigorously evaluated throughout model development, and have incorporated the feedback of the Peer Review Panel. Further discussion of the spatial and temporal needs is provided in Section 4 of the FMD.

- Errors in programming, preprocessing, running, linking, and post-processing the models—A thorough evaluation of model validity includes the stage of model verification (EPA, 1994; Schwartz, 2000; Shelly et al., 2000), which ensures that mathematical relationships posited by the model are correctly translated into computer code, that model inputs are free of numerical errors, and that software and data quality assurance has been conducted. EPA conducted such model verification throughout the modeling study.

- Untested model predictions—EPA acknowledges that there are some model predictions that cannot be evaluated using model-data comparisons because field data are inadequate or lacking. However, this is not a “new” or additional source of uncertainty because the component uncertainties related to the model predictions have already been acknowledged.

Considerable effort went into the analysis of model uncertainty in this project, which is perhaps unprecedented for a suite of water quality models of this complexity. The MVR evaluates the uncertainty of the models using approaches based on Monte Carlo analysis. This approach basically assumes that the models are correct, and errors in model predictions arise from uncertainty or variability in parameter values. These are estimated based on subjective expert judgment using probability distributions, and parameters are generally assumed to be uncorrelated. While these assumptions are obviously suspect, they are nonetheless common practice in uncertainty analysis. Analyzing the uncertainty of the linked model predictions is a strength of this work.

Response 5B-DE-2:

EPA agrees with the Reviewer's comments. Although some of the assumptions required for the Monte Carlo approach rely on professional judgment, EPA concurs that such simplifications were necessary to conduct an uncertainty assessment that was computationally feasible, and that such methods were consistent with the state of the practice in complex environmental fate and transport models.

The response surface modeling of EFDC uncertainty appeared to be very thorough and comprehensive. EPA has provided a descriptive field for the parameters listed in Table 5.2-8, which hopefully will be incorporated into the report. I would like to see EPA add initial and boundary conditions for PCBs to this analysis, since these are both uncertain factors. I disagree that the RSM approach is impractical or unworkable. The Responsiveness Document (p.2-110 and 111) discusses how uncertainty analysis can be used in context of evaluating differences between remediation scenario predictions. Hopefully, no more than one RSM will be necessary for evaluating uncertainty of these predictions.

Response 5B-DE-3:

The uncertainty approach presented in the MVR was intended to be demonstrative rather than prescriptive, and EPA recognizes that other approaches to evaluating model uncertainty may also be valid. EPA appreciates the Reviewer's recognition of the value of the uncertainty analysis; however, in subsequent discussions with the Model Working Group following the Peer Review of the MVR, it was agreed that the uncertainty of the model results in the evaluation of the relative performance of remedial alternatives would be addressed using the best professional judgment of the Model Working Group.

The Monte Carlo analysis of the FCM is also quite thorough and informative. I agree with the suggestion to evaluate model uncertainty in terms of 10th and 90th percentiles, as being reasonably conservative, consistent with the distribution properties of the food chain model outputs. This is probably the simplest way to deal with inflation of predictive uncertainty. The critique of "unreasonable" parameter sets producing model predictions outside the bounds of the data, is really a criticism of the subjective estimation of parameter variability. This can be addressed using a Bayesian Monte Carlo approach, which builds posterior (informed) parameter distributions based on computing the likelihood function for each model realization[1]. The method is simple and works very well - although the computational burden of BMC (i.e., effort to achieve convergence in the likelihood function) can be heavy, unless there is some way to substitute the RSM for the numerical model. I don't know whether this is a practical option for computationally-intensive models such as EFDC.

Response 5B-DE-4:

EPA agrees with the Reviewer's comments. EPA believes that the computational burden of a Bayesian Monte Carlo approach outweighs the potential benefits of the more sophisticated approach.

There has also been much discussion about how residuals from comparisons between model predictions and data can be used to improve the uncertainty analysis. It has been suggested to use residuals to directly characterize uncertainty. This strikes me as a naïve approach, as I would normally consider residual error to be a "best case" estimate of uncertainty. Forecasts will seemingly always be more uncertain than calibration; the question is, how much more? It has also been suggested that reporting average bias (e.g., Tables 4.2-4, 4.2-10 and 4.3-2) masks the uncertainty due to variability in the PCB data itself. The remedy to this is straightforward: calculate and report the variance of model bias. Moreover, statistical model could be developed from the residuals to estimate the distribution of PCB concentrations from mean predictions. This would compliment, instead of replace, the Monte Carlo results. The difficulty I can foresee, is how do you determine the proper assumptions regarding homoscedacity (e.g., is the variance in a constant factor of the mean or simply a constant?).

Response 5B-DE-5:

EPA agrees with the Reviewer's comments. Comparisons of model predictions to data are useful, but do not provide an indication of overall model uncertainty, particularly when the model is to be extrapolated to new exposure scenarios. The reporting of the variation in the model bias metric had been added to the FMD (Section 3.8); therefore, this gap in reporting has already been addressed.

EPA has not conducted statistical modeling to estimate concentrations in individual fish (i.e., using modeled averages and estimating variation based on observed variability in field data). Although the suggested approach is reasonable, the model objectives do not require the estimation of PCB concentrations in individual organisms, but instead rely on comparisons of absolute and relative concentrations of average PCB concentrations for alternative remedial scenarios.

---

[1] Dilks, D.W., Canale, R.P. and P.G. Meier. 1992. Development of Bayesian Monte Carlo techniques for model quality model uncertainty. Ecological Modeling. 62:149-162.

Marcelo H. Garcia

While uncertainty in model inputs and outputs was recognized, the nature of the model does not allow for a conventional uncertainty analysis. Once all the processes are accounted for and well represented in the model and assuming that the EFDC code can be run more efficiently, it might be feasible to conduct an uncertainty analysis.

Response 5B-MG-1:

The Reviewer's comment does not address model uncertainty but instead repeats concerns discussed under General Topic 1, which relates to the balance of fate processes. EPA disagrees that questions regarding the suitability of model process formulations preclude assessment of the model uncertainty. Because models are iterative in nature, the results of sensitivity and uncertainty assessments are commonly used to guide model refinements. It is not useful or feasible to wait for consensus on all aspects of model development prior to conducting or interpreting a model uncertainty assessment.

Frank Gobas

The calculation of model uncertainty in the report is seriously flawed. The report includes several references to this. For example, based on findings that the uncertainty in the calculated PCB concentrations is greater than the variability in the sampling data, the authors conclude that "the results should not be interpreted to mean that the uncertainty in model predictions renders the model predictions too uncertain to be usable" (p.5-56 Model Validation Report). The authors further "acknowledge that a true statistical analysis of uncertainty, particularly when uncertainty is propagated through the modeling framework, can produce bounds that may not be possible (or likely) based on an understanding of that system. (p.5-57 Model Validation Report). I think that these are important points.

Response 5B-FG-1:

EPA disagrees strongly with the Reviewer's comments and believes that the Reviewer has misinterpreted the comments made by EPA in the MVR. The comments by EPA regarding the differences between the width of the concentration distributions in the data and those of the model simulations remain appropriate, and are by no means an admission of a "flawed" approach. To the contrary, EPA's comment was intended to provide a clear statement that the uncertainty analysis includes quantification of uncertainty from all sources, not just uncertainty in input data, and therefore, if conducted properly, can never result in bounds that are smaller than the inherent variability in the data. The Reviewer's comments are diametrically opposed to those of another Reviewer on this topic (see Response 5B-DE-2).

I agree that in models of this complexity, it is virtually impossible to meet some key criteria for a meaningful Monte Carlo Simulation, namely (i) that the model parameters included in the simulations are independent, and (ii) there is insufficient data to properly characterize the variability/uncertainty/error in the many parameters used in the model. Also, it should be recognized that the MCS method does not recognize error in the functional relationships of the model. The KS method is subject to the same limitations as the MCS method and the failure of 4 out of the 56 runs may be due to implausible parameter selections causing the model to crash for these runs due to the problems outlined.

In conclusion, the conditions for a meaningful MCS or KS uncertainty analysis are not met in this study. The calculated uncertainty values are therefore seriously flawed. It is unclear what the meaning of the calculated uncertainty is. The lack of a meaningful uncertainty analysis is a major flaw of the current model because model projections are very difficult, if not impossible to interpret.

Response 5B-FG-2:

EPA agrees that, in the strictest sense, it is not possible to meet some of the statistical assumptions of the MCS approach (i.e., complete independence of model inputs). However, EPA believes that the simplifications inherent in the approach were necessary and do not unduly compromise the utility of the linked model uncertainty assessment. EPA’s opinion is substantiated by the comments of another Reviewer on this topic (see 5B-DE-2), who indicated that, in spite of the inability to fully satisfy the strict conditions of MCS, the overall approach was “common practice” and “a strength of this work.”

EPA acknowledges that professional judgment was sometimes required in the establishment of input parameter distributions, and that the MCS approach does not explicitly incorporate error in the functional form of model processes. However, EPA does not believe that the conditions for a meaningful MCS or KS analysis are not met and that the calculated uncertainty values are thus seriously flawed. Further, as stated in the comments of 5B-DE-2, the MCS method can be considered along with other types of uncertainty assessment. Finally, the MCS approach used in the modeling study is consistent with published model applications at other sites in the northern United States (Burkhard, 1998).

To include uncertainty and better characterize it, the modelers could consider using the formidable empirical data set to calculate frequency distributions for the model predicted mean concentrations. To formalize this method, the authors could further develop the MB method described on p. 6-118 (Model Validation Report) by calculating the standard deviation of the mean (i.e. MB). What this will do in the case of PCB concentrations in the sediments is simply project the observed variation in PCB concentrations on the model predicted concentrations. The result is now a distribution of predicted concentration that is grounded in empirical data. This is not a major job, and could be done with little extra work.

Response 5B-FG-3:

EPA agrees that further assessment of the model-data comparisons is useful, and has further developed the MB method as suggested by the Reviewer. Results of the expanded analysis are presented in Section 3.8.3 of the FMD.

Recommendation:

My recommendation is to ignore the MCS and KS results and remove it from the model framework when the model is to be applied in the next phase of the study. Instead I recommend that the modelers use the discrepancy between model predicted concentrations and observed concentrations as a measure of model uncertainty. This is simpler, easier to understand and avoids current computational problems. For example, the data depicted in Figures 6.3.3 to 6.3.8 (Model Validation Report) can provide a reasonable description of the overall model uncertainty. This can be achieved by calculating the confidence limits of the MB used in the report. See Environ. Toxicol. Chem. 23, 2343-2355 for additional details on this method. The application of this method will also generate frequency distributions of model outcomes that can be used in risk assessments. The method that I recommend is not complicated and can be carried out in little time. The main drawback of the application of this method is that it relies on the assumption that the uncertainty identified in past application of the model is a good measure for uncertainty in future model applications. I think that this is a reasonable assumption for some model applications. However, if river functioning is drastically altered by the remediation efforts, this assumption may not apply.

Response 5B-FG-4:

The Reviewer's proposed solution entails discarding all explicit assessment of parameter uncertainty, and relying solely upon model-data comparisons. This approach, if used in isolation, would be flawed for three main reasons:

- The alternative approach ignores many aspects of model uncertainty, and in fact ignores more aspects of uncertainty than the MCS approach. Simple model-data comparisons do not explicitly account for uncertainty in model formulations, parameter uncertainty, analytical uncertainty in the data, or aggregation errors. To this end, EPA concurs with the opinion of another Reviewer (5B-DE-2) that reliance solely on model residuals is a "naïve approach."
- The alternative approach cannot assess objectively whether the model has been overcalibrated. Assessment of magnitudes of model residuals does not indicate whether these residuals can be expected to be stable when the model is applied to new conditions. Taken to an extreme, the model could be artificially forced to perfect numerical fit (i.e., zero bias and zero residuals) by inappropriately manipulating model parameters, but the model-data approach would not diagnose this problem. This is because all sources of uncertainty are buried in a "black-box" using this method, such that sources of deviation are not discernible from one another.
- The alternative approach relies on the assumption that the magnitude of uncertainty identified during Calibration will remain unchanged in future model applications. This assumption may not hold if the model is applied to new exposure conditions in remedial scenarios. The Reviewer acknowledged this limitation.

EPA believes that ignoring the MCS and KS results would negatively influence, rather than benefit, the uncertainty analysis. However, the "observed discrepancy" approach suggested by the Reviewer has some merit, and for this reason, an expanded assessment of the MB approach is included in the uncertainty analysis in the FMD. The inability to strictly adhere to all assumptions is a limitation of any uncertainty analysis method, including the MCS and KS approaches and the Reviewer's recommended alternative.

Any model has inherent uncertainty. So, this is normal. But where the model framework is inadequate is in the recognition of this uncertainty and in the development of adequate methods to characterize or estimate this uncertainty. Without the inclusion of uncertainty in the current model framework, I do not think that it is possible to convincingly distinguish between the effectiveness of different remedial options.

Response 5B-FG-5:

EPA strongly disagrees with the Reviewer's comments. EPA has acknowledged uncertainties of many types, and was able to explicitly account for some, but not all, of these sources of uncertainty during the linked MCS approach. It is inappropriate to characterize the EPA uncertainty assessment as non-existent or inadequate when the alternative recommended approach suffers from some of the same, plus several additional, limitations.

EPA believes a more productive approach is to acknowledge that the methods can be considered in a complementary manner to evaluate model uncertainty. As noted in the FMD (Section 3.8), methods for the quantitative evaluation of uncertainty in modeling study results are in various stages of development and there is no commonly accepted way of quantifying or communicating to decisionmakers the accuracy of model predictions. By considering these approaches together, EPA believes that a meaningful and defensible evaluation of model uncertainty was conducted.

Include model uncertainty in the model framework and provide guidance about how the results from the uncertainty analysis are to be used when comparing outcomes resulting from different model scenarios.

Do not use the MCS or KS method for calculating model uncertainty.

Include an uncertainty analysis that takes full advantage of the empirical data that have been collected. As discussed earlier, this can be achieved by comparing observed and model predicted PCB concentrations.

Response 5B-FG-6:

Model uncertainty was included in the model framework, and has been explored in additional detail in the FMD based on Reviewer suggestions, including additional comparisons of observed and model-predicted PCB concentrations.

In subsequent discussions with the Model Working Group following the Peer Review of the Model Validation Report, it was agreed that the uncertainty of the model results in the evaluation of the relative performance of remedial alternatives would be addressed using the best professional judgment of the Model Working Group rather than using a formal uncertainty analysis as proposed in the MVR.

Wilbert Lick

Because the basic processes are inaccurately described, the uncertainties in model outputs are not correctly acknowledged or described.

Response 5B-WL-1:

The Reviewer comment does not address model uncertainty but instead repeats concerns discussed under General Topic 1, which relates to model processes. EPA disagrees that differences of opinion regarding the suitability of model process formulations should obstruct all assessment of the model uncertainty. See also the Response 5B-MG-1 above.

E. John List

Since the basic processes for the fate and transport of PCB are clearly not properly included in the model their uncertainty cannot be assessed.

Response 5B-JL-1:

The Reviewer comment does not address model uncertainty but instead repeats concerns discussed under General Topic 1, which relates to model processes. EPA disagrees that differences of opinion regarding the suitability of model process formulations should obstruct all assessment of the model uncertainty. See also the Response 5B-MG-1 above.

6. GENERAL TOPIC 6: UPSTREAM BOUNDARY CONDITIONS

6.1 COMMENT SUMMARY

Reviewer Name	Page Number/Line Number
Adams	Page 5, lines 39 to 45
	Page 7, lines 1 to 23
Endicott	Page 4, line 33 to page 5, line 22
	Page 6, lines 5 to 16
	Page 10, line 18 to page 16, line 9
Gobas	Page 6, lines 34 to 40
	Page 7, lines 35 to 36

6.2 GENERAL TOPIC SUMMARY

PCB transport and fate in the Housatonic River is strongly influenced by the concentrations of solids and PCBs that enter the study area from the East Branch of the river near the GE facility.

Three of the Reviewers commented, one extensively, on the relationships developed to provide upstream boundary input of TSS and PCBs based on flow and stage. All three Reviewers noted that the relationship overpredicted TSS loadings during low-flow periods and mentioned that this issue affects the parameterization of the model throughout the PSA. One Reviewer noted that the model results appear extremely sensitive to upstream boundary conditions and questioned whether in-stream sediment processes were important in the model. Two Reviewers recommended additional data collection, and all three recommended additional analysis of the TSS and PCB data to refine the boundary loadings.

One Reviewer focused on the change in PCB boundary conditions between the Calibration and Validation phases of the modeling study and requested additional explanation of and justification for the changes. The same Reviewer made a number of suggestions for refined analyses of the boundary condition data, including use of the Autobeale statistical approach.

6.3 GENERAL TOPIC RESPONSE

The upstream boundary at the confluence of the East and West Branches of the Housatonic River is the largest source of solids and PCBs to the study area. Solids and PCB loads from the upstream boundary must be estimated as accurately as practical to ensure model reliability. Throughout the model development process, flow, solids, and PCB concentrations and other measurements were repeatedly examined to ensure that upstream loads were estimated in the most realistic manner possible. Because of their relative importance to PCB transport, particular attention was paid to methods used to estimate upstream boundary loads for the East Branch. The load estimation methods used were refined during the modeling study, reflecting the improved understanding of the system, as well as in response to Peer Review comments.

Solids and PCB loads are computed by multiplying paired values of concentration and flow together in sequence over time. To develop inputs for EFDC, point-in-time measurements of concentration must be extended to form a continuous record and multiplied by flow of upstream loads to the river for the entire time period simulated by the model. The flux analyses described in Attachment B.3 of the MCR provides an overview of data and the initial methods used to estimate upstream boundary solids loads. Further data explorations and analysis are presented in FMD Appendix B.2.

Data collected during storm event and monthly monitoring programs were used for these analyses. It should be noted that solids and PCB concentrations measured at the upstream boundary are highly variable. Across narrow ranges of flow, these measured values can vary by more than two orders of magnitude. For all practical purposes, a many-to-one relationship exists between measured concentration and flow. This variability complicates load estimation efforts because load estimation techniques such as regression estimators typically describe the central tendency (e.g., the mean or median) of the relationship between variables such that there is a one-to-one relationship between concentration and flow. When data are highly variable, there can be considerable differences between the central tendency (the one value estimated for a given flow) and any individual measurement. Extensive efforts were directed toward determination of optimal methods for load estimation, assessment of variability, and an evaluation of the significance of differences between boundary concentration estimates and measurements.

As specifically recommended during Peer Review, the Autobeale method (Richards, 1998) to estimate loads as a function of concentration and flow was evaluated. The Autobeale method uses a sequence of concentration and flow estimates to compute a load based on the assumption that the ratio between load and flow is constant for a set of conditions. The program can segregate (stratify) the data into different flow regimes as needed to report the mean daily load for 1 year of data. The program contains an algorithm to identify the optimal stratification based on the smallest pooled mean square error.

Despite the conceptual appeal of this method, the Autobeale approach cannot be used to estimate loads for EFDC because the program is designed to report an average daily load for a 1-year period. This is a serious limitation because loads with high temporal resolution must be specified for EFDC to generate the high-resolution model outputs required as a primary objective of the modeling study. For example, loads for EFDC must be estimated with a frequency of 15 minutes to 1 hour as needed to simulate high-flow events. Further, the Autobeale program cannot be used to estimate loads for years for which there are no concentration measurements. Moreover, when there are few measurements in a year, the program typically assumes that concentration is constant for long periods of time regardless of the flow record and the expectation that concentration will increase substantially as flows increase. Because it cannot estimate loads for years without concentration observations, the Autobeale program cannot estimate loads for the majority of the years simulated as part of the modeling study (i.e., loads cannot be estimated for most of either the model calibration or validation periods).

Many other approaches for estimating TSS and PCB boundary loads were evaluated. Relationships between TSS and PCB concentrations and flow, flow acceleration, rate of acceleration, and precipitation were explored. None of these approaches produced better estimates of TSS loads than the method first used for model calibration (see the flux analysis presented in Appendix B.3 to the MCR). However, the stratified regression estimation procedure described in the Section 2.3 of Appendix B.2 to the FMD resulted in a revised estimate of PCB boundary conditions which yields estimates that fit the data somewhat better than prior estimation methods did.

One issue extensively examined as part of model development efforts was the nature of the error in boundary condition estimates. For this discussion, two terms must be defined: (1) residual error; and (2) bias. Residual error is defined as a difference between the value of an estimate and the value of a measurement. Bias is defined by the sign (positive or negative) and consistency of residual errors. A positive (or high) bias occurs when residual errors (residuals) are consistently greater than zero and a negative (low) bias occurs when residuals are consistently less than zero.

In an ideal case, boundary condition estimates would exactly reproduce each and every measurement so that the residual error for each estimate would be zero. However, because concentration measurements for the Housatonic River have a wide range of values across narrow ranges of flow, residual errors occur. Stochastic and deterministic methods were extensively explored as part of efforts to minimize residual errors and reduce the impact of bias in upstream boundary concentration estimates. However, neither the stochastic nor deterministic bias correction methods were successful at reducing bias without increasing residual error.

6.4 RESPONSE TO SPECIFIC COMMENTS ON GENERAL TOPIC 6

Eric Adams

Boundary Loads

The trend of the PCB concentration versus flow data for the E. Branch, shown in Figure 4.1-2 seems strange (why the sudden change at 550 cfs?). I guess this is simply what the data suggest. Also the model fit exceeds the data at low flow (10 cfs) by a factor of 2-3 for both dissolved and particulate PCBs. This excess upstream load may have led to an underestimation of calibrated fluxes from other sources downstream. Given the prevalence of low flow periods, this is important.

Response 6-EA-1:

The change at 550 cfs is related to the flow at which out-of-bank flow conditions occur in Reaches 5 and 6. As noted by the Reviewer, the increase in PCB concentrations when flows exceed 550 cfs is what the data suggest. Differences between model results and data at low flows are driven by the very wide range of measured PCB concentrations that occurs. At these flows, PCB concentrations are typically low but vary from non-detect to 40 mg/kg. TSS also varies over a wide range at any given flow. The regression estimators for the upstream load were formulated to accurately estimate the central tendency (mean) of the measurements at any flow rate for both TSS and PCBs. PCB loads are estimated as a function of the TSS concentration to reduce the potential for bias in the model.

However, it should be recognized that a factor of 2-3 difference in TSS and PCB concentrations at low flow (which is when low concentrations occur) results in a small difference between simulated and measured concentrations because both concentrations are low (i.e., the relative difference is small in an absolute sense because measured concentrations are low). Further, EPA took great care in establishing the TSS and PCB boundary conditions in EFDC to ensure that any potential biases in the boundary conditions did not translate to errors in food chain exposures.

Also, see the General Response above, the General Response to General Topic 1B, and Appendix B.2 and Section 4 of the FMD.

TSS

The model uses flux analysis for upstream TSS when there is no data and often this analysis overestimates TSS by a factor of 2-4 (Figures 4.2-31 and 6.2-12). There is some underestimate of TSS at high flow.

The model tends to overpredict TSS during the calibration period (Figure 4.2-39), especially at lower TSS concentrations. This is possibly because of the overprediction of upstream loading. The model seems to do better with the storm events, which EPA claims are more important. While bias in TSS at low flow may be insignificant in terms of the sediment budget, it does affect the water column PCB concentrations, and hence the relative uptake of PCBs from the water column and the sediment in the Food Chain Model. In general it would have been nice to have had more data on TSS.

The statistical summary (equivalence plot, Figure 4.2-61) indicates generally good agreement, but the model shows less variability than the data; i.e., the model overpredicts low concentrations and underpredicts high concentrations. There appears to be no significant bias with flow rate. Sedimentation rates agree reasonably well with measurements using Cs-137.

Bedload concentration is computed as bed load mass rate divided by river flow (pg. 4-56). Given that the bedload is traveling slower than the average stream velocity, this would seemingly give too low concentrations, but I suspect velocities are more important than concentrations.

Response 6-EA-2:

See Response 6-EA-1, the General Response above, and Appendix B.2 to the FMD.

Douglas Endicott

This question also asks us to consider changes to the models resulting from additional (Phase 2) calibration. This specifically applies to EFDC calibration, since little or no changes were made to either the HSPF or FCN models. Previously, the Peer Review panel commented that the 14-month calibration was too short to reflect many significant long-term model processes, specifically the evolution of PCB concentrations in the sediment bed. So the EPA modeling team went back and extended the calibration to cover a much longer, 10-year duration. I expected to see some changes in the model parameters related to the residence time of PCBs in the surficial sediment layer. Instead, as a result of Phase 2 calibration we see some dramatic changes in the simulation of PCB transport and fate in the Housatonic River: There are now 3 times more PCBs entering the Primary Study Area (PSA) at the upstream boundary, although the difference (compared to the Phase 1 calibration & flux analysis) diminishes further downstream (2 times more PCBs at New Lenox Road, 1.5 times more at Woods Pond). Export of in-place PCBs from the sediments of Reach 5A is no longer the predominant contaminant source to the Rest-of-River. To me, this is a breathtaking change in the PCB fate and transport simulation, which rewrites the entire assessment going back all the way to the conceptual model presented in the MFD. I was disturbed to see such a major change taking place so late in the model development process, and was also somewhat surprised at the nonchalance with which this revision was presented and received.

Does the hydrograph or the longer duration of the Phase 2 calibration explain these differences? To some degree, maybe. In 10-year period, there are a number of high-flow years (e.g., ‘90, ’96,’98 and ‘00) as well as low-flow years (’91, ’92 and ‘99). The hydrograph impacts the balance of sediment erosion and deposition averaged over different durations. The MVR does not discuss this to any significant extent. However, the change in the PCB fluxes mostly reflects a new treatment of the boundary condition, apparently provoked by some new boundary condition data. Comments regarding the boundary condition used in the Phase 2 calibration are provided in my response to question 3. At this point, EPA should consider whether further refinements and improvements to the model will occur in the future, as more data are collected and the models are run. In other words, are the models only as good as the latest data point? How can the model calibration be considered “final” if, as the Phase 2 recalibration demonstrates, a few new data can so readily change the PCB flux analysis and possibly other model results?

Response 6-DE-1:

As the Reviewer notes, some differences between the Phase 1 and Phase 2 Calibrations results are attributable to the differences in flow over time. As the Reviewer also comments, some of the difference is attributable to the changes in upstream boundary PCB loads using the revised method used to estimate loads, particularly at high flow. In the Housatonic River, measured sorbed PCB

> concentrations (in mg PCB/kg solids) increase as TSS concentrations increase when flows exceed 550 cfs. The PCB load function first developed for the MCR (and used for the Phase 1 calibration) did not capture this trend in the measurements. Using the original functions for TSS and PCB loads, the sorbed PCB concentration estimated at the upstream boundary would decrease rather than increase as measured. This condition was not noted during Phase 1 model development because the range of flows over the 14-month period does not include many high-flow events. However, when the original method was used to estimate PCB loads for the Phase 2 calibration period, the dilution of sorbed PCB concentrations at the upstream boundary that was introduced by the original load estimation functions became apparent because of the numerous high-flow periods that occur from 1990 through 2000.
>
> Many methods to estimate loads were evaluated to address this limitation. The best method to accurately simulate the increasing relationship between sorbed PCB concentrations, TSS, and flow was to estimate PCB loads as a function of TSS concentration and flow. This approach ensures that PCB concentrations at the upstream boundary match observed trends at high flow and also minimizes the potential for bias in water column biotic exposure concentrations at low flow.
>
> EPA disagrees with the Reviewer's assessment that the difference in simulated PCB fate is "breathtaking." In light of the extensive analyses devoted to this issue, EPA further disagrees with the Reviewer's characterization that revisions to PCB upstream boundary conditions were introduced with "nonchalance." It is important to recognize that all other parameters (mass transfer coefficients, erosion and deposition parameters, etc.) are unchanged. It should also be noted that the model representation of erosion and deposition or other processes is not directly tied to the method used to estimate upstream boundary loads. This means the balance of fate processes is correctly represented and that the model can be considered "final" because the model is able to accurately simulate water column TSS and PCB concentrations, net sedimentations rates, and the trend of surface sediment PCB concentrations in Woods Pond over time. Additional data for the upstream boundary will not alter the parameterization of PCB transport and fate processes in the model domain.
>
> Further, it is also important to note that remediation efforts are being conducted on the GE facility and nearby areas, including the 2 miles of the East Branch between the facility and the confluence. These remediation efforts have altered PCB loads at the upstream boundary. Consequently, the PCB upstream boundary condition used for Calibration and Validation is not applicable for the future application of the model in the CMS and must reflect the loads that are currently measured or expected to occur in the future. It is expected that this will be a topic of discussion for the Model Working Group.

EFDC predictions of flow, TSS and water column total PCB concentrations were compared to data using both longitudinal profile and time series plots. Although the spatial profiles for TSS predictions appear reasonable for several selected days (Figures 4.2-34 thru 4.2-37), the time series plots indicate significant and persistent overprediction of low TSS measurements. This is particularly evident at Holmes Road (Figure 4.2-38), but also at New Lenox Road (4.2-39) and to a lesser degree at Woods Pond headwater (4.2-40). In the MVR (page 4-59), EPA offers that "Simulated TSS concentrations pass through approximately one-third of the data…", which is surprisingly poor model performance. Similar overprediction of low TSS measurements is apparent for the validation period as well. According to comments by both EPA and GE, the overprediction of low TSS measurements is due to significant bias in the specification of the East Branch boundary condition on days when TSS concentrations were not measured. Time series plots for specific flow events (Figures 4.2-42 thru 4.2-48) indicate a better fit of TSS predictions under high flow conditions.

Response 6-DE-2:

The greatest difficulty associated with TSS and PCB load estimation for the Housatonic River is the very wide range of measured concentrations that occur across narrow ranges of flow. As noted in Appendix B.2 to the FMD and other documents, TSS and PCB concentrations are particularly variable at low flow, when concentrations are typically lower. Natural variability reflects the diversity of watershed conditions. Because of this variability, a many-to-one relationship exists between measured concentration and flow. This complicates load estimation efforts because regression estimators describe the central tendency of the relationship between variables so there is a one-to-one relationship between concentration and flow.

Given this variability, differences between boundary concentration estimates and individual measurements occur and are unavoidable (i.e., the average is neither the minimum nor the maximum). However, it should be recognized that small errors in simulated TSS or PCB concentrations that occur under low-flow conditions have relatively little impact on model performance in any context because the mass flux (i.e., the product of flow and concentration) of solids and/or PCBs is at a minimum when flows are low.

Further, EPA took great care to ensure that any potential biases in model results were minimized and did not influence the EFDC or FCM calibration process. Moreover, the endpoint of primary concern with respect to model bias is the ability of the model framework to estimate PCB exposure concentrations and body burdens. In this regard, there is no significant bias in model performance, as demonstrated by the bounding analysis presented in the EPA Response to Questions from the Model Validation Document Overview Meeting and further described in FMD Section 4.

Therefore, EPA disagrees that there is a "serious overprediction" of low TSS and PCB concentrations in EFDC as described by the Reviewer. Also see General Response above.

The first two issues appear related to parameterization errors, as noted previously, and should be satisfactorily addressed by the suggested corrections. The bias in TSS and PCB concentrations due to the boundary condition deserves further consideration.

At sampling locations in the upper reaches of the PSA, the model fails to match half to two-thirds of the TSS data. Predicted values at base flows are ~10 mg/L versus data in the 2-5 mg/L range. There is less discrepancy in the lower reaches. In general, the bias is most apparent at lower flow rates. Censoring TSS data <5 mg/L was offered by EPA as a way of correcting the model performance metrics. A similar problem with low-flow bias for PCB predictions is also apparent in the upper reaches of the PSA, and a similar censoring approach was suggested. However, I do not believe it is appropriate to ignore or discount the bias in model predictions under low-flow conditions. About 70% of an “average year” is low flow conditions, and significant PCB uptake by fish coincides with low flow periods. It is unclear whether the low-flow PCB bias has led to food chain miscalibration, as suggested by GE, because the food chain model was initially calibrated using data-based exposure concentrations. These would not exhibit the same low-flow bias as exposure concentrations predicted by the model. Nevertheless, I am concerned that bias due to the TSS and PCB boundary conditions may result in model errors that have not been addressed. We know very little about how the boundary condition will evolve over time following upstream remediation; hopefully, ongoing monitoring will be used to update and improve upon this situation.

The specification of the upstream PCB boundary condition is an important component of developing a reliable transport and fate model for the Housatonic River, for each of the time intervals over which the model is to be applied (calibration, validation and forecasting). EPA has developed different interpretations of the PCB boundary condition in the Phase 1 and Phase 2 Calibrations, and these differences produce dramatically different outcomes in terms of the PCB flux analysis generated by the model predictions. The PCB boundary condition developed for the Phase 2 calibration applies the partitioning model to estimate particulate and dissolved PCB concentrations from total PCB measurements, which are then independently regressed to flow rate. The PCB phase concentrations are “capped” to prevent the boundary conditions from grossly exceeding the data at high flows. This is a creative approach; I tried something similar in the Fox River. Unfortunately, it’s not clear whether this approach does a reasonable job of describing the boundary condition data. Although this approach may improve the PCB boundary condition at high flows, in comparison to the Phase 1 calibration, at base flow rates the Phase 2 PCB boundary conditions are substantially worse (e.g., Phase 1: ~40 ng/L at base flow vs. Phase 2 : ~100 ng/L). I am not convinced EPA has made the best use of the boundary condition data, and offer some comments below regarding the approach they used to describe this data. Regardless of these issues, it appears that much of the variability seen in PCB concentrations at the upstream boundary cannot be explained, in the approaches taken in either the Phase 1 or Phase 2 calibration, or other alternatives that have probably already been explored. If the TSS and PCB boundary conditions cannot be improved, the best recommendation may be to add a random error component to the boundary conditions in the sensitivity and uncertainty analyses.

The boundary condition data eres provided by EPA in the spreadsheet “East_Branch_Boundary_Data.xls”. Their interpretation of this data, and the descriptive approach used to model the boundary condition, is described in Section 4.1.5 of the Model Validation Report, with further discussion offered in supplemental responses. Figure 4.1-2 (reproduced below) shows the regressions of dissolved and particulate PCB concentrations versus flow that were used to specify the total PCB boundary condition for the EFDC transport and fate model (except for days when measurements were available). I have a few comments to make based on this figure and review of the boundary condition spreadsheet:

1. As a general observation, I would say that if one focuses on the actual data plotted in these graphs (the open circles), it appears that EPA's regressions describe relatively little of the variability observed in PCB concentrations at the boundary condition. The regressions are especially tenuous at low and high flows. In fact, the regressions exceed <u>all</u> of the data in the lowest flow "bin".

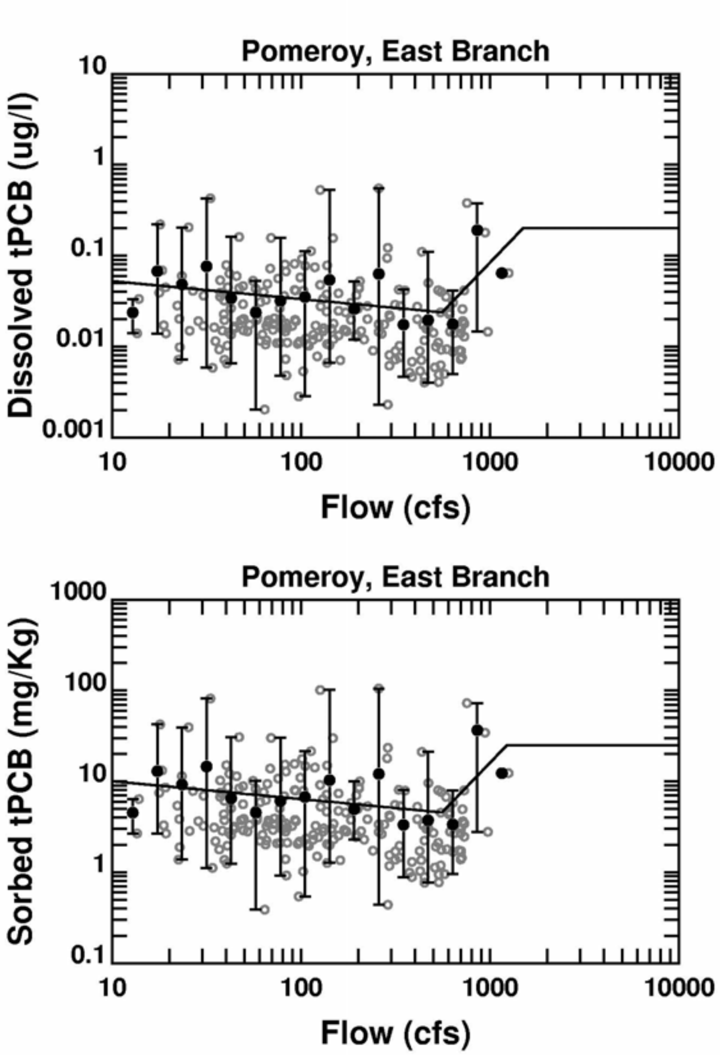
2. I cannot find the PCB concentration data collected at flow rates smaller than 24 cfs[1]. Therefore, I have no way to confirm the results shown in Figure 4.1-2 for the two lowest-flow "bins". At least for the lowest "bin", the regression line appears to be a poor fit. A noted above, it exceeds all of the measured data.

3. At high flows, the PCB concentrations are "capped" to prevent unrealistically high PCB concentrations during floods. The "cap" was specified as 25 mg/Kg for particulate PCBs, based upon concentrations measured in sediments deposited on vegetation and banks. I think this value may be somewhat too high. My own calculation of the unbiased mean PCB concentration from the 13 sediment deposit samples is 20.5 mg/Kg, although I was measuring the data off of a graph. Please check this calculation. If 25 mg/Kg is indeed too high, it will bias the high-flow PCB boundary condition.

4. According to EPA, the detection limit was used as replacement values for censored PCB concentrations. While this has little impact on the distribution of total PCB concentrations, it does bias the particulate PCB concentrations that are calculated from total PCB measurements using the partitioning model. I didn't check, but I suspect that dissolved PCB concentrations may be biased as well. I think it is generally preferable to use one-half of the detection limit as replacement values for censored PCB concentrations.

---

[1] I think this may be because East Branch daily flows were provided in the spreadsheet, while Coltsville (hourly?) gauge data were used in the regressions and possibly the graph. I couldn't find the source of the flow data in the MVR.



5. Another thing I noticed when I reproduced EPA's partitioning calculation (which is used to estimate the dissolved and particulate PCB concentrations from the total PCB measurements), is that the dissolved PCB concentration estimates don't agree very well with the relatively few dissolved PCB concentrations measured at the boundary condition. This is shown below in Figure 2. If the partitioning model is being used to describe the boundary condition, shouldn't it fit the data?

6. Frankly, some of this data is mystifying. For example, 2 samples were collected on November 12, 1995. For both samples, the flow is reported to be over 1,100 cfs. In the first sample, [TSS] was 450 mg/L and total [PCB] was 0.2 ug/L. In the second sample, [TSS] was 17 mg/L and total [PCB] was 0.52 ug/L. Running these values through the partitioning model, one obtains particulate PCB concentrations of 0.44 and 23 mg/Kg. It is somewhat hard to believe that PCB concentrations on suspended solids could vary by a factor of 50 within one day.

7. It is also curious to me that there has been no shift or evolution in PCB boundary concentrations over the past 16 years, and that seasonality is not evident in the data. If PCB boundary condition data for all years are plotted together, we see the familiar shotgun target (Figure 3). There appears to be some underlying behavior, however, within individual years of data. Figure 4 shows total PCB data for years (especially 2002) in which a positive relationship with flow is evident. Figure 5 shows total PCB data for years (2003 in particular) in which the flow relationship is generally negative. The highest total PCB concentrations in Figure 5 come from the warm (and presumably higher organic carbon[2]) months of July-September, while the lowest PCB concentrations at very high flow were sampled in October and November. While I am confident EPA has looked at this data in great detail, I think the patterns in the data hinted at here suggest some underlying behavior which could be exploited in the regression models used to describe the boundary condition. I believe this might, in turn, provide better PCB boundary condition estimates than the current approach in which all measurements are lumped into one sample.

8. The MVR doesn't discuss what other alternatives may have been tested, but here are some suggestions:

   - AutoBeale (Beal's Stratified Ratio Estimator) is an alternative statistical approach that works well for boundary conditions when the data are stratified by season as well as flow;

   - When regressing the PCB data, were autocorrelation and Lag-1 or 2 considered as factors?

   - Check bias correction formula;

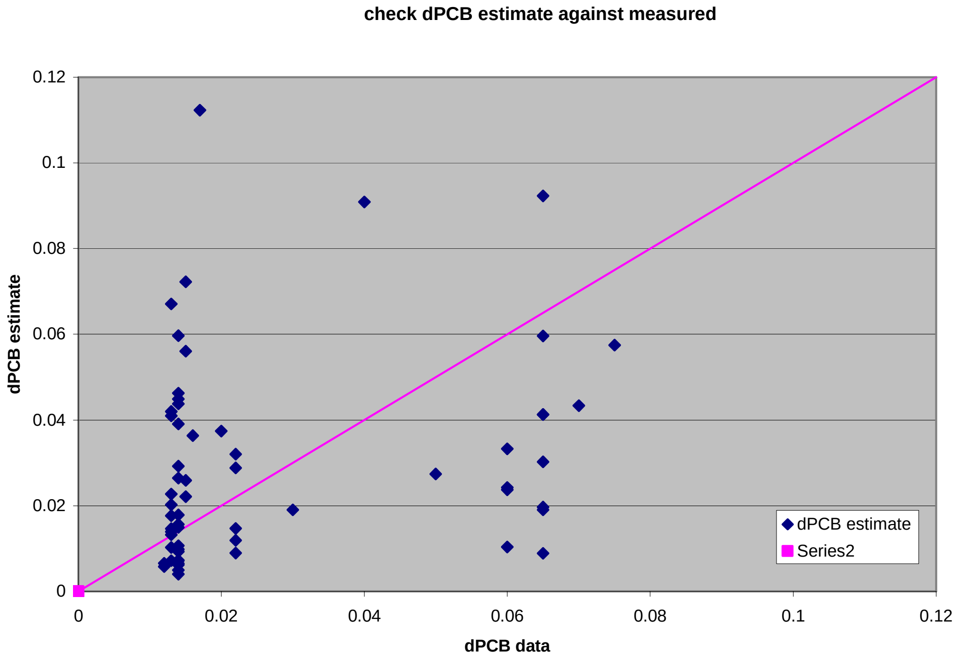
   - Have untransformed PCB data been regressed? How do results compare to log-transformed regressions?

---

[2] EPA is assuming a constant 7.3% organic carbon content on suspended solids, which cannot be confirmed from the data collected at the upstream boundary.

1

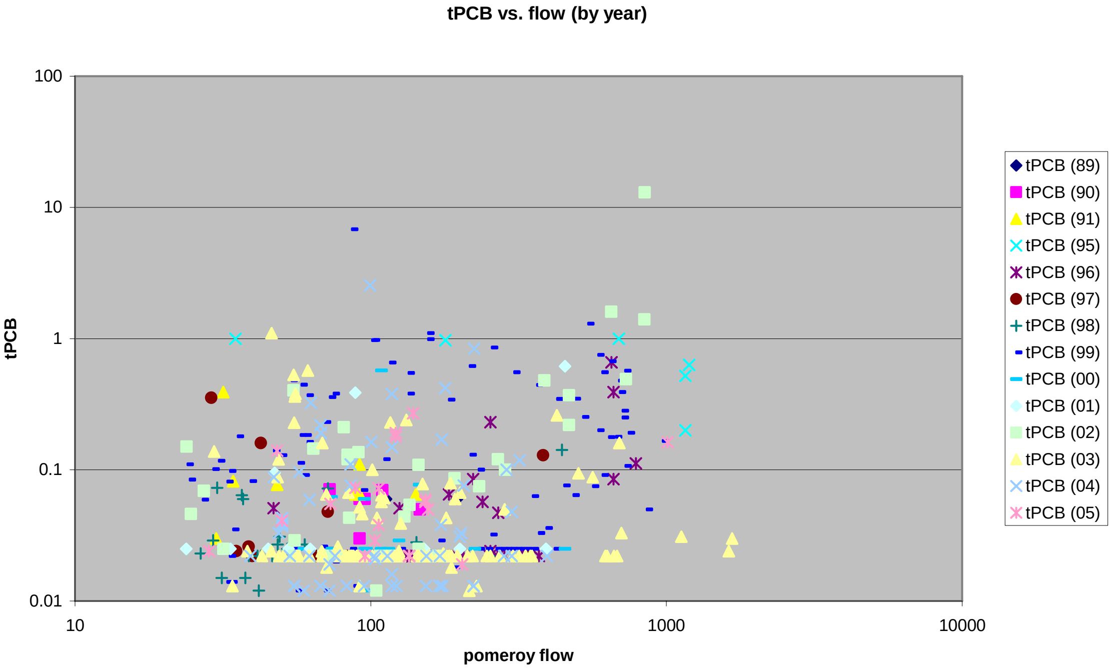
Figure 2



2

3

Figure 3



4

5

Figure 4

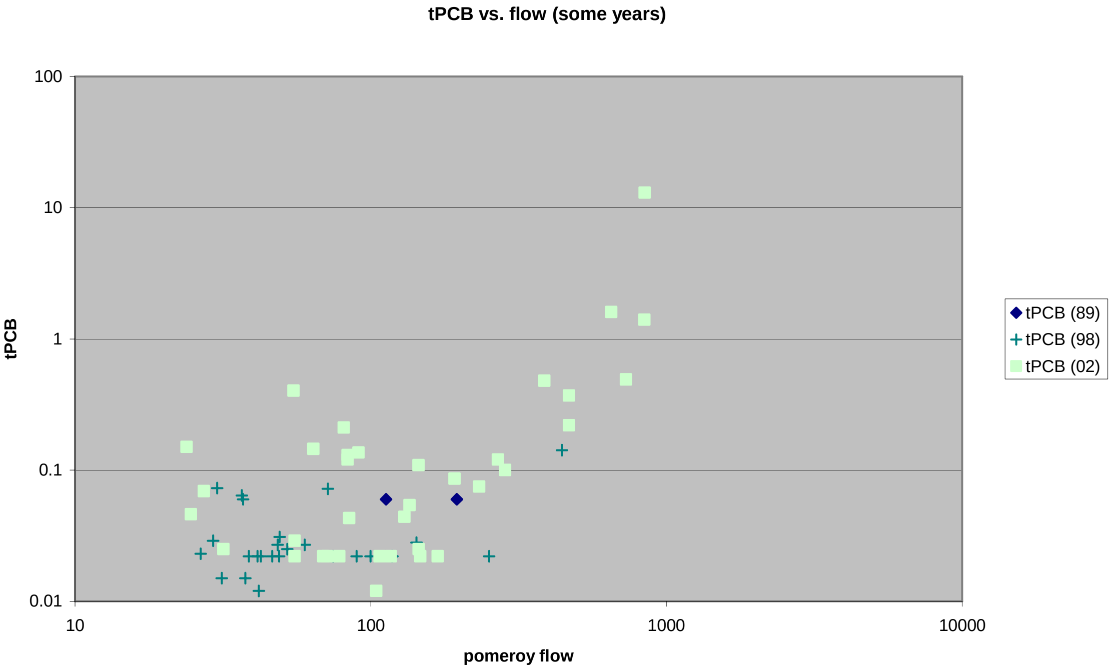
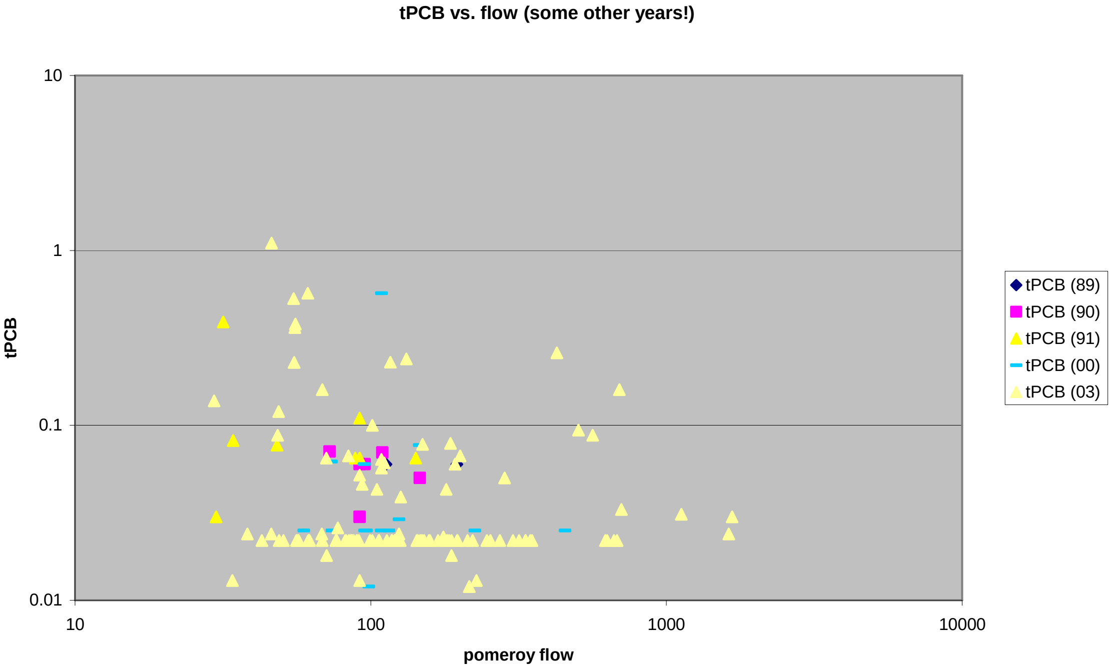
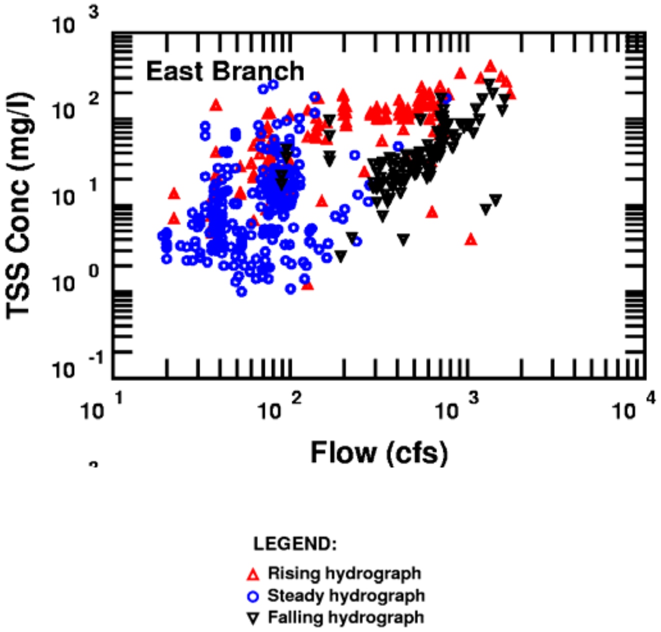


Figure 5



9. I also went back to look at the TSS boundary condition, which has not been changed from the Phase 1 calibration. Figure 6 is a reproduction of the plot displaying the East Branch TSS data as a function of flow (from: Volume 3—Appendix B Hydrodynamic and Sediment/PCB Fate and Transport Model Calibration, Attachment B.3 TSS and PCB Flux Analysis). The TSS data were partitioned according to 3 phases of the hydrograph – rising, steady and falling. TSS data are well-behaved during rising and falling hydrographs, but not at steady (and predominantly low) flow. This makes me wonder again about seasonal factors, intermittent upstream disturbances, or possibly sampling artifacts. It is hard to imagine TSS concentrations varying between 1 and 100 mg/L at base flow.

Figure 6



Response 6-DE-3:

EPA appreciates the detailed and thoughtful comments provided by the Reviewer regarding suggested load estimation techniques and approaches. As documented in FMD Appendix B.2, as well as in the MCR and MVR, many load estimation methods were explored, including many of the Reviewer suggestions. In particular, the use of the Autobeale approach was evaluated and stochastic and deterministic methods to reduce differences between boundary concentration estimates and individual observations were explored.

The greatest difficulty associated with TSS and PCB load estimation for the Housatonic River is the very wide range of measured concentrations that occur across narrow ranges of flow. As noted in Appendix B.2 to the FMD and other documents, TSS and PCB concentrations are particularly variable at low flow, when concentrations are typically lower. This natural variability reflects the diversity of conditions in the watershed. As a result of this variability, a many-to-one relationship exists between measured concentration (or load) and flow for all practical purposes. This variability complicates load estimation efforts because load estimation techniques such as regression estimators typically describe the central tendency (e.g., the mean or median) of the relationship between variables such that there is a one-to-one relationship between concentration and flow.

As a result of variability, differences between boundary concentration estimates and individual measurements occur and are unavoidable. However, EPA took great care to ensure that potential biases in model results were minimized and did not influence either the EFDC or FCM calibration process. In particular, FCM results are not biased as a result of PCB upstream boundary condition estimates.

It is worth noting that in several instances the Reviewer's comments focus on particulate PCB concentration (expressed as μg PCB/L). When TSS and PCB boundary conditions are considered jointly as the sorbed PCB concentration (expressed as mg PCB/kg of solids), there is little bias in the PCB boundary condition. This is the reason why there is little bias in FCM exposure results.

For further details, see the General Response above, FMD Appendix B.2, and FMD Section 4 (FCM bounding analysis), as well as the MCR and MVR.

Frank Gobas

First, in the majority of comparisons between measured and model predicted TSS concentrations, the TSS concentrations are over predicted. This does not appear to have a corresponding impact on the calculation of the PCB concentrations in water. The effect of this bias on the calculation of PCB concentrations under current conditions appears to be low. However, the bias may become important when the model is applied under different conditions. It is therefore important to explore the reasons for the bias in TSS and make appropriate corrections.

Response 6-FG-1:

See General Response above as well as Sections 3 and 4 and Appendix B.2 of the FMD. As previously noted, TSS and PCB concentrations are particularly variable at low flow, when concentrations are typically lower. As a result of this variability, a many-to-one relationship exists between measured concentration (or load) and flow for all practical purposes. This variability complicates load estimation efforts because load estimation techniques such as regression estimators typically describe the central tendency of the relationship between variables such that there is a one-to-one relationship between concentration and flow.

As noted by the Reviewer, this does not have a significant impact on simulated PCB concentrations in the water column. This is because the PCB boundary condition estimate accurately simulates sorbed PCB concentrations (mg PCB/kg solids). Further, the parameterization of the model framework was specifically developed to ensure that potential biases in one framework component would not adversely impact subsequent results in any other component. In particular, TSS boundary conditions for EFDC were extensively evaluated with respect to bias and the impact that bias has on FCM results. It should be noted that the apparent bias in EFDC results for low-flow conditions is almost entirely attributable to the natural variability in TSS concentrations at the upstream boundary.

I recommend that the modelers explore model parameterization schemes that remove bias in the TSS model predictions.

Response 6-FG-2:

See General Response above as well as Sections 3 and 4 and Appendix B.2 of the FMD. See also Response 6-FG-1.

7. GENERAL TOPIC 7: SEDIMENT INITIAL CONDITIONS

7.1 COMMENT SUMMARY

Reviewer Name	Page Number/Line Number
Adams	Page 6, lines 1 to 5
Bohlen	Page 10, lines 15 to 24
Endicott	Page 8, lines 5 to 8

7.2 GENERAL TOPIC SUMMARY

Finite difference models require the assignment of concentrations in each computational grid cell at the beginning of each simulation; these are referred to as initial conditions. The application of EFDC to the Housatonic River includes water column, bed sediment, and floodplain cells. Specification of initial conditions is required in each type of cell.

Three Reviewers commented on the procedure used in model Validation to establish initial conditions for sediment PCB concentrations. All three noted that the hindcasting approach did not allow for a robust independent test of model performance, but differed on an appropriate response.

7.3 GENERAL TOPIC RESPONSE

The majority of the data for PCB concentrations in environmental media were collected in 1998-1999, at the end of the Phase 2 Calibration period. In conducting the 10½-year Phase 2 Calibration from January 1990 to June 2000, it was recognized that use of the initial sediment PCB concentrations from the short-term Phase 1 Calibration (i.e., projecting the conditions measured in 1998-1999 to the January 1, 1990 model simulation start date), would result in initial conditions that could underestimate the PCB concentrations in sediment in Reaches 5 and 6. There were insufficient data to establish initial conditions using a data-based approach; therefore, an alternate procedure was necessary. The method used to establish initial conditions was based on using the rate of decline in PCB concentrations in surface sediment simulated by the model as a hindcast to estimate PCB concentrations at the January 1, 1990 Phase 2

Calibration start date (see Section 4.1.7 of the MVR). These annualized rates of attenuation were used to calculate scaling factors that were applied to the 1998-1999 data to adjust the initial PCB conditions in each model grid cell at the start of the Phase 2 calibration and validation periods, respectively.

The single exception to this procedure was for Woods Pond (Reach 6) for the validation simulation, where adequate data were available from sampling conducted in 1980 to allow the determination of initial sediment PCB concentrations from data. The consistency between the initial condition of 975 mg/kg OC for Woods Pond, based on 1980 data, and the estimate of 999 mg/kg OC developed by applying the scaling factor procedure used for the other reaches, provides evidence that the scaling factor approach is reasonable. This approach is consistent with other components of model parameterization, where site-specific data are used when available.

7.4 RESPONSE TO SPECIFIC COMMENTS ON GENERAL TOPIC 7

Eric Adams

Initial conditions for sediment PCB concentrations

I agree with GE that hindcasting predicted trends is not a robust way to establish initial conditions. By playing the tape backwards, then forwards, they will get the same conditions they started with. I am not sure what the alternative is, but the lack of independence should be acknowledged.

Response 7-EA-1:

EPA agrees that hindcasting predicted trends is not a robust way to establish initial conditions. The Reviewer's statement that "I'm not sure what the alternative is…" reflects the assessment of the modeling team that led to the approach that was implemented. EPA acknowledged in MVR Section 4.1 (Changes to EFDC Model Application) that this approach was adopted because of the lack of historical data to set initial conditions. The discussion of temporal changes in simulated PCB concentrations in bed sediment (MVR Section 6.2.3.2.3) states:

> *Data collected during 1980 were used to establish initial conditions in Woods Pond for the validation simulation. In other reaches, data were not available to set initial conditions for the validation simulation, and instead, the model was used to estimate initial conditions. Therefore, Woods Pond represents the only*

portion of the PSA where sufficient historical data exist for a direct assessment of simulated temporal changes in sediment PCB concentrations.

In both the MVR and in discussions with the Peer Review Panel during the MVR Document Overview Meeting, EPA attempted to make it clear that limited historical data complicated the assessment of temporal trends simulated by the model, in both the floodplain and in the river sediments. The one exception was in Woods Pond, where historical data were available to develop initial conditions for PCBs in the sediment. In that case, the estimated concentration of 999 mg/kg OC, developed by applying hindcasting approach, was consistent with the initial condition of 975 mg/kg OC derived directly from the 1980 data. This comparison provided a degree of confidence in the approach.

W. Frank Bohlen

With regard to PCBs the data set allows at least an initial evaluation of the spatial and temporal validity of the model. Specification of the initial concentrations used in the Validation remains a concern since to some extent the hind cast method used to develop initial concentrations is not entirely independent of the subsequent model run. This apparently was required by the limited data available for 1979. Review of the RCRA Facility Investigation Report (BBL, 2003) shows the results of sampling dating back to the 1979-1980 period. It seems possible to use this variety of data to provide at least a check on the trends and associated initial concentrations suggested by use of the model. Might this be possible? I'm assuming that extensive "data mining" has been part of this exercise and that therefore some of this approach might already have been tried. If so, a brief discussion of this in the report would be useful.

Response 7-FB-1:

The Reviewer is correct in concluding that the limited historical data was the reason for the use of the hindcast method to develop initial conditions for 1979 (with the exception of Woods Pond, where sufficient data were available, and were used). In Reaches 5A through 5C, the number of sediment samples collected in 1980 averaged approximately 12 samples per mile. Given the variability in PCB concentrations in Housatonic River sediment, which can be two orders of magnitude within 1 m^2, this data density was considered far too sparse to use to make reliable estimates of initial conditions.

In response to Reviewer's comments, additional descriptions of the data and methods used to develop initial conditions for the sediment and floodplain soil have been compiled in Appendix B.3 to the FMD.

Douglas Endicott

Furthermore, I agree that the approach for developing initial conditions for sediment PCBs (circa 1990 for the Phase 2 calibration and 1979 for the validation) precludes a "robust test" of the long-term model predictions. Consequently, these initial conditions should become part of sensitivity and uncertainty analyses.

Response 7-DE-1:

The sediment initial conditions were not added to the sensitivity analyses, which involved a uniform adjustment of ±50% of the base case value, because this level of change is not realistic based on available data, especially fish tissue data.

8. GENERAL TOPIC 8: UPSTREAM MODEL

8.1 COMMENT SUMMARY

Reviewer Name	Page Number/Line Number
Adams	Page 8, lines 12 to 35
Bohlen	Page 12, lines 1 to 3
Endicott	Page 2, lines 23 to 24
	Page 5, lines 23 to 30
	Page 9, lines 29 to 32
	Page 18, lines 24 to 25
	Page 21, line 37
	Page 24, lines 7 to 21

8.2 GENERAL TOPIC SUMMARY

During Validation, the upstream boundary of the EFDC model for the East Branch was re-evaluated to address the morphological changes following remediation of the 2.4 kilometers (1.5 miles, Reaches 3 and 4) immediately upstream of the confluence, which was completed in spring 2006. The purpose of this exercise was to construct and test a revised upstream boundary model that would be representative of the upstream area for use in future-casts of the model to evaluate baseline conditions and potential remedial alternatives. The transport and fate of PCBs were not simulated in this model application because of the scarcity of applicable data from the downstream portions of Reach 4 due to the ongoing remediation over the past 6 years. The upstream model simulations were run using the hydrograph from the 25-year model validation period (excluding 2004).

Three Reviewers did not consider the upstream model to be adequately calibrated/validated and commented that insufficient data were available to evaluate the model performance, or recommended that more data be gathered or work be performed and/or more discussion of the model performance be provided. In addition, it was mentioned that no sensitivity or uncertainty analyses or scenario projections were performed for this component of the model framework.

8.3 GENERAL TOPIC RESPONSE

The modeling team has further evaluated the necessity for the application of the upstream model and discussed it with the Model Working Group, including representatives from GE and their consultants. The decision has been made not to pursue the use of the upstream model. The collective belief is that application of the upstream model within the modeling framework will not make a significant difference in the output from the model simulations. Therefore, it is not necessary to use the upstream model to achieve the goal of the modeling study to evaluate baseline conditions and potential remedial alternatives.

8.4 RESPONSE TO SPECIFIC COMMENTS ON GENERAL TOPIC 8

Eric Adams

Upstream Model

The model was extended two miles upstream to include the 0.5 and 1.5 mile remediation areas. Upstream Q and TSS were obtained from rating curves, while upstream tPCB = 0.019 microgram/L from boundary measurement. Comparisons were attempted against TSS and tPCB measured at Lyman St. in 2001-2004 (some of this is reported in DRM handouts and not in the MVR).

Predicted TSS varied from ~$3x10^{-4}$ to >10^{2} mg/L while data were in the range of 3 to 30 mg/L (mostly 3-10). [I am assuming these are the correct units for TSS; the vertical axis on the handout distributed at the DRM says ng/L.] Can one really predict TSS as low as $3x10^{-4}$ mg/L? Some data are in the range of the model, but model predictions show much more variability? Where does this come from?

Predicted PCBs were $3x10^{-5}$ to 10^{-3} mg/L while data were in the range of ND up to $3x10^{-5}$. [Again, I am assuming these are the correct units; they are consistent with the 30 to 1000 ng/L indicated on the handout GE provided, though EPA's handout during the document review meeting indicates $3x10^{-5}$ to 10^{-3} ng/L.] In any case, the agreement is not very convincing, and ***it is hard to say that the upstream model is validated***.

The downstream concentrations of TSS and PCB from the upstream model are much lower than the corresponding upstream boundary values used for the main model. Given that output from the main model depends so heavily on its upstream boundary condition, more work should be done on the upstream model, or at least more interpretation of the results. There is scant discussion in the MVR.

Response 8-EA-1:

See General Response above.

W. Frank Bohlen

The above comments refer only to the PSA. It would appear that there are insufficient data to adequately test model results in both the upstream (which is in a state of major change) and the downstream model. Efforts to gather these data should be initiated immediately.

Response 8-FB-1:

See General Response above.

Douglas Endicott

I should note here that, since the upstream and downstream models share the same frameworks, the same comments apply.

Response 8-DE-1:

See General Response above.

Another change that has been undertaken by EPA is to extend the spatial domain of the models to cover the upstream reach of the East Branch through the area that has been remediated, as well as downstream from Woods Pond dam to Rising Pond dam. It is encouraging to see this development, because the modeling domain confined to the ROR was clearly too small to fully address the extent of PCB contamination in the Housatonic River. However, I am not convinced that either the upstream or downstream models are adequately calibrated and validated at this time. It is obvious that much less effort has gone into these models, and the presentations made by the EPA modeling team at the Document Review Meeting looked much more like progress reports than finished products.

Response 8-DE-2:

See General Response above.

In the upstream model, there are too few TSS data to tell whether the model predictions are reasonable or not. For both TSS and PCBs, my impression is that the model predictions are much more variable than the few available data. PCB concentrations appear to be substantially overpredicted (e.g., a factor of 5 to 10 overprediction in PCB concentrations).

Response 8-DE-3:

See General Response above.

No sensitivity or uncertainty analyses have been presented for the upstream and downstream models.

Response 8-DE-4:

See General Response above.

No scenario projections were presented for either the upstream or downstream model.

Response 8-DE-5:

See General Response above.

I do not believe that either the upstream or downstream models are sufficiently developed to be adequate to address the goals of the modeling study. Their development appears to have been rushed, an impression reinforced by the fact that they apparently did not exist a year ago. Sensitivity and uncertainty of the upstream and downstream models have not been reported, and documentation of the development of these models has been inadequate. In its present status, the upstream model is incomplete and not fully validated. PCB fate and transport have not been simulated, and insufficient model-data comparisons have been made and/or reported. Incorporation of the upstream model in the MVR appears premature, given that substantial additional development efforts will be required before it can be used reliably. Likewise, the downstream model has not been sufficiently validated. Limited model-data comparisons suggest there may be significant bias in PCB concentration predictions. There also appears to be a lack of adequate and appropriate data for this significant extension of the model domain. In their current states of development, neither the upstream nor downstream models are suitable for use in modeling future conditions in the Corrective Measures Study. I am hopeful that the inadequacies of the upstream and downstream models can be addressed through continuing monitoring and modeling activities.

Response 8-DE-5:

See General Response above.

9. GENERAL TOPIC 9: MODEL APPLICATION TO REACHES 7 AND 8 (DOWNSTREAM MODEL)

9.1 COMMENT SUMMARY

Reviewer Name	Page Number/Line Number
Adams	Page 8, line 37 to page 9, line 13
Bohlen	Page 12, lines 1 to 3
Endicott	Page 2, lines 23 to 24
	Page 5, lines 23 to 30
	Page 9, line 32 to page 10, line 4
	Page 18, lines 24 to 25
	Page 21, line 37
	Page 24, lines 7 to 21

9.2 GENERAL TOPIC SUMMARY

The model domain for the Phase 1 Calibration included Reaches 5 and 6, the 10 ½-miles from the confluence of the East and West Branches to Woods Pond. Comments were received in previous Peer Reviews from the Reviewers and GE suggesting that the model domain be extended both downstream and upstream. EPA responded by extending the model domain downstream for EFDC and FCM to Rising Pond Dam (Reaches 7 and 8), an additional 19 miles. HSPF already included this portion of the watershed in the Calibration and Validation; therefore, no changes to the HSPF domain were necessary.

Three Reviewers commented directly on the need for further work on the downstream model. These Reviewers commented on the comparative lack of data from the downstream reaches, with two of these Reviewers recommending additional data collection efforts. Two Reviewers noted that there was some general correspondence between the limited TSS and PCB concentration data and concentrations predicted by the downstream model, but that there were insufficient model-data comparisons to demonstrate that the downstream model was sufficiently calibrated and/or validated for use in predicting PCB fate and transport following potential remedial alternatives.

9.3 GENERAL TOPIC RESPONSE

EPA recognized the relative scarcity of data useable for model development downstream when originally making the decision to include only Reaches 5 and 6 in the EFDC and FCM domain. In response to Reviewer comments, EPA expanded the model domain downstream to include Reaches 7 and 8 following the MCR Peer Review. Cross-sectional surveys of the river channel were performed to establish channel bathymetry, and field surveys were conducted to verify the resulting elevation data, floodplain conditions, and aquatic habitat. This information supported the development of a model grid and initial assignment of parameters for EFDC for Reaches 7 and 8 (see Section 3.5 of the FMD and Section 6.4 of the MVR). It also provided the information necessary to subdivide Reach 7 by hydrologic regime and aquatic communities. These subreaches were then assigned fish community types comparable to the corresponding subreach in Reaches 5 and 6 for FCM simulations (see Section 3.6 of the FMD and Section 6.4 of the MVR).

In response to Reviewer comments in this Peer Review, the model Calibration for Reaches 7 and 8 was revisited to be consistent with the changes made to the calibration parameters in Reaches 5 and 6. The time period for application of the model was extended from 2 years (1997–1998) to a 15-year simulation (1990–2004). In addition, changes were made to the parameterization of physical conditions to better represent the very different hydrologic regimes (e.g., steep gradients, larger grain sizes) that occur in Reach 7.

The model is driven entirely by the results from other models (i.e., flows and sediment and PCB loads from Reach 6 [Woods Pond] simulated by EFDC, and tributary and nonpoint source flows and loads simulated by HSPF). This factor, and the application of the calibrated models for Reaches 5 and 6 to very different physical conditions and PCB concentrations in Reaches 7 and 8, provides a demonstration of the robustness of the EFDC and FCM applications for the Housatonic River. Model-data comparisons show that the models performed adequately in this extended application. The final model validation for EFDC and FCM in Reaches 7 and 8 is presented in Section 3.5 and 3.6, respectively, of the FMD.

9.4 REPONSE TO SPECIFIC COMMENTS ON GENERAL TOPIC 9

Eric Adams

Downstream Model

The model was extended 19 miles downstream to Rising Pond Dam (not described in MVR; only in DRM handouts). Upstream boundary conditions were taken from output of the main model at Woods Pond. Model-data comparisons for TSS and tPCB were made between 1990 and 2004 ~ 1-mile downstream from Rising Pond Dam.

Predicted TSS varies widely as do the data, with the peaks corresponding with peaks in flow. They are in the same ballpark, but the variability is so great that you need a statistical comparison to determine how well they match on average.

Predicted PCB concentrations also vary substantially (~$3x10^{-3}$ to ~3 microgram/L), with predicted peak concentrations corresponding with peak TSS and peak flow, and non-peak concentrations hovering generally around 0.03 micrograms/L. (But why are some up to 10 times lower?) Measurements show much less variability with an average of around 0.1 microgram/L, or about 3 times the predicted non-peak values. The data are too sparse to see if they respond to peaks in flow. Again a statistical comparison would be helpful to quantify model-data agreement. Both measurements and predictions seem to be 2-3 times lower than corresponding values at Woods Pond Outlet, which makes sense given the 20 mile separation.

GE's plot of simulated versus measured PCBs at Rising Pond Dam shows very little correlation between simulation and measurement ($r^2 = 0.01$). ***At this point it is hard to say the downstream model is validated***.

Response 9-EA-1:

In response to the Reviewers' comments, additional work on the downstream model was performed following the Peer Review of the MVR as described above. Model-data comparisons for flow and TSS, and PCB concentrations demonstrate that EFDC adequately simulates these parameters (see Section 3.5 of the FMD). The differences shown in the cross-plots of the simulated versus measured TSS and PCB concentrations and the residual and relative residual plots in Section 3.5 of the FMD are attributed in large part to phasing differences between the simulated results at Rising Pond Dam and the data collected from Division Street. In addition, phasing and volumetric differences between actual and simulated flows from HSPF for runoff and tributaries in Reach 7 contribute to the phasing difference between measured and simulated PCB and TSS concentrations. One example of this occurs when flows from HSPF are higher than the actual flows, resulting in lower simulated PCB concentrations than measured concentrations due to dilution.

The simulated water column and sediment PCB concentrations also result in good agreement between the measured fish tissue PCB concentrations and those simulated by FCM in the linked application. This is another line of evidence for the adequacy of the downstream model validation and in general for the robustness of the EFDC parameterization.

W. Frank Bohlen

The above comments refer only to the PSA. It would appear that there are insufficient data to adequately test model results in both the upstream (which is in a state of major change) and the downstream model. Efforts to gather these data should be initiated immediately.

Response 9-FB-1:

EPA does not believe that it is necessary to collect more data to demonstrate the adequacy of the model for Reaches 7 and 8 for the reasons stated above.

Douglas Endicott

I should note here that, since the upstream and downstream models share the same frameworks, the same comments apply.

Response 9-DE-1:

The Reviewer comments that the application of EFDC to the upstream and downstream models shares the same frameworks as those used for Reaches 5 and 6, thus his comments regarding the application for Reaches 5 and 6 apply. The Reviewer’s comments on the Reach 5 and 6 application of EFDC are addressed in other sections of this Responsiveness Summary.

Another change that has been undertaken by EPA is to extend the spatial domain of the models to cover the upstream reach of the East Branch through the area that has been remediated, as well as downstream from Woods Pond dam to Rising Pond dam. It is encouraging to see this development, because the modeling domain confined to the ROR was clearly too small to fully address the extent of PCB contamination in the Housatonic River. However, I am not convinced that either the upstream or downstream models are adequately calibrated and validated at this time. It is obvious that much less effort has gone into these models, and the presentations made by the EPA modeling team at the Document Review Meeting looked much more like progress reports than finished products.

Response 9-DE-2:

See the General Response and additional responses provided above.

In the downstream model, predicted TSS and water column PCB concentrations are in the same ranges as the data at the Rising Pond outlet, but there is not much fidelity beyond that. Initial downstream model PCB predictions included in the MVR clearly indicated a problem with the balance between erosion and deposition in high-gradient reaches. As presented at the Document

Review meeting, this has been addressed by adding an additional noncohesive sediment class in the sediment transport simulation.

FCM predictions in the downstream model domain are compared to PCB concentrations measured in fish for Reach 7 and Rising Pond. These predictions look favorable overall, although there may be bias in the predictions for specific species. Results appear comparable in data-based and linked exposure predictions.

Response 9-DE-3:

The model calibration was revisited to address the patterns of erosion and deposition noted by the Reviewer, and subsequent validation was performed following the Model Validation Peer Review Meeting.

EPA agrees that the model bias statistics for the downstream model are “favorable overall.” The Model Bias (MB*) statistic tracks the central tendency of the ability of the model to simulate PCB concentrations by category. The overall MB* for the downstream model (linked to EFDC predictions) is 1.08, meaning that on average the model is overpredicting PCB concentrations by 8%.

With regard to possible bias for specific FCM species, Table 3.6-11 in the FMD shows MB* broken down by species for the linked application of FCM. Largemouth bass, bluegill, yellow perch, and pumpkinseed samples show no consistent bias. Brown bullhead are somewhat overpredicted, by 44% on average. Bluntnose minnows are systematically overpredicted; however, this bias statistic is based on five samples. In addition, bluntnose minnows are not modeled explicitly in FCM and, therefore, require extrapolation from a representative species (dace).

No sensitivity or uncertainty analyses have been presented for the upstream and downstream models.

Response 9-DE-4:

A sensitivity analysis for the application of EFDC to Reaches 7 and 8 was performed and the results are presented in Section 3.7 of the FMD. The sensitivity analysis for FCM (Section 3.7.3 of the FMD) was conducted using the model application for Reaches 5 and 6. This model parameterization was remained unchanged when applied to Reaches 7 and 8. Therefore, there was no need to repeat the sensitivity analysis. The model sensitivities in Reach 8, if calculated, would be virtually identical to those for the Reach 6 sensitivity simulations. The Reach 7 sensitivities would be equivalent to the model sensitivities for the corresponding subreaches in Reach 5.

No uncertainty analyses were performed for the downstream model. See General Topic 5B.

No scenario projections were presented for either the upstream or downstream model.

Response 9-DE-5:

The Reviewer is correct, the example scenarios were simulated for Reaches 5 and 6.

I do not believe that either the upstream or downstream models are sufficiently developed to be adequate to address the goals of the modeling study. Their development appears to have been rushed, an impression reinforced by the fact that they apparently did not exist a year ago. Sensitivity and uncertainty of the upstream and downstream models have not been reported, and documentation of the development of these models has been inadequate. In its present status, the upstream model is incomplete and not fully validated. PCB fate and transport have not been simulated, and insufficient model-data comparisons have been made and/or reported. Incorporation of the upstream model in the MVR appears premature, given that substantial additional development efforts will be required before it can be used reliably. Likewise, the downstream model has not been sufficiently validated. Limited model-data comparisons suggest there may be significant bias in PCB concentration predictions. There also appears to be a lack of adequate and appropriate data for this significant extension of the model domain. In their current states of development, neither the upstream nor downstream models are suitable for use in modeling future conditions in the Corrective Measures Study. I am hopeful that the inadequacies of the upstream and downstream models can be addressed through continuing monitoring and modeling activities.

Response 9-DE-6:

EPA agrees that the downstream model benefited from further development and subsequent Calibration and Validation following the Peer Review (see Sections 3.5 and 3.6 of the FMD), and believes that it is now adequate for use in evaluating the relative performance of potential remedial alternatives. It was recognized at the outset of model development that the data available for calibration and validation of any model of downstream reaches were more limited than the data available for Reaches 5 and 6. Additional information was collected to allow the application of EFDC and FCM to Reaches 7 and 8. One of the purposes of modeling Reaches 7 and 8 was to test the robustness of the EFDC model as calibrated for Reaches 5 and 6. As noted in presentations to the Peer Review Panel and in the FMD, the parameterization of the EFDC model used in Reaches 5 and 6 was changed slightly (e.g., the addition of the fourth non-cohesive sediment class) for the downstream model during the calibration process.

Statistical evaluation of EFDC model-data comparisons for TSS and PCB concentrations indicate that the relative bias at Rising Pond Outlet is well within the model performance measure of ± 30% for TSS (-11.93%) and just outside the measure for PCB concentrations (-31.97%). For median relative error, the model performance measure is also ±30%, and the EFDC model is within the performance measure for both TSS (-27.12%) and PCB (-3.32%) concentrations.

Model performance, as quantified by the relative bias and the median relative error, and by the ability of the linked application of FCM to simulate PCB concentrations measured in fish tissue, is judged to be adequate to achieve the goal of the modeling study.

10. GENERAL TOPIC 10: EXAMPLE SCENARIOS

10.1 COMMENT SUMMARY

Reviewer Name	Page Number/Line Number
Adams	Page 11, lines 23 to 41
	Page 12, lines 4 to 8
Bohlen	Page 14, lines 6 to 16
Endicott	Page 19, line 5 to page 20, line 32
	Page 21, lines 4 to 36
Garcia	Page 6, lines 26 to 28
Gobas	Page 12, lines 7 to 45
Lick	Page 2, lines 1 to 3
List	Page 5, lines 5 to 25
	Page 6, lines 2 to 15

10.2 GENERAL TOPIC SUMMARY

Upon request from a Reviewer during the Calibration Peer Review, a qualitative assessment of model performance was conducted by examining model simulations using the 26-year validation period hydrograph. These two hypothetical scenarios were run altering the initial conditions similar to what might be done when using the model in the Corrective Measures Study. The first example scenario was simulated with a zero PCB load at the East Branch upstream boundary. The second example scenario was also simulated with a zero PCB load from the East Branch upstream boundary, but in addition included zero PCB concentrations in Reach 5A sediment, and no PCB or solids loadings from the banks in Reach 5A.

The input parameters for each hypothetical scenario were adjusted to reflect a change in river conditions, and the simulation was allowed to proceed without further adjustment. Changes in final PCB concentrations in the main channel sediment and floodplain soil resulting from each simulation were qualitatively reviewed for reasonableness of model behavior.

Four Reviewers commented that the model response from the two scenarios appeared reasonable in the sense that most locations experienced some decline in PCB concentrations in surficial sediment. A Reviewer considered the response in PCB concentrations in the water column and fish tissue concentrations simulated in the example scenarios to be reasonable and plausible.

In addition, single Reviewers commented on very specific aspects of the simulations. These comments and responses are listed in Section 10.4.

10.3 GENERAL TOPIC RESPONSE

EPA agrees with the Reviewer comments that the model results from the two scenarios are reasonable. Changes in final PCB concentrations in the main channel sediment and floodplain soil resulting from each simulation were qualitatively reviewed for reasonableness of model behavior. Within this context, the model reasonably reflects system-level responses that might be anticipated when loads are reduced or eliminated. The pattern of changes in bed sediment PCB concentrations is consistent with known properties and dynamics of sediment transport in the river. The nature and magnitude of the reductions of PCB concentrations in the surficial floodplain soil are consistent with the expectation that the input of clean solids and subsequent transport onto the floodplain during out-of-bank flood events would reduce PCB concentrations over time. The spatial pattern of these simulated reductions is generally consistent with the locations of known high-deposition areas on the floodplain. EPA believes these simulations provided an additional test of model performance.

It is important to recognize that the scenarios simulated in this exercise are not intended to reflect actual, proposed, or anticipated remedial alternatives, and no further significance should be attached either to the hypothetical scenarios evaluated or the simulation results. Under the terms of the Consent Decree, baseline conditions and potential remedial alternatives are to be identified and evaluated as part of the Corrective Measures Proposal and Study to be conducted by GE.

10.4 RESPONSE TO SPECIFIC COMMENTS ON GENERAL TOPIC 10

Eric Adams

I appreciate the inclusion of the two example simulations that shed some light on how the model might perform when simulating remediation options. This is clearly a start. There would seem to be two issues to consider. The first is how well the remediation measures themselves will work. For instance, will a cap erode? The second is what will happen downstream (in space and time) after remediation? For example, will dirty sediments cover up capped or dredged sediments? The example simulations don't consider the first question (though it is very important), but partially address the second. However, several considerations come to mind. First, the example simulations were begun with "cold start" initial sediment conditions, which seems to explain, at least in part, why sediment concentrations increased in some areas downstream from simulated remediation (i.e., places where there was no PCB source). Better that the model be spun up to initial sediment conditions that reflect an equilibrium with the modeled hydrology and bathymetry. Also, the model seems to ignore any particulate PCBs that might have wanted to deposit in the channel of Reach 5A. (What happens to these in the model?) Ultimately, the model could be generalized to look at "different colored" PCBs emanating from different source locations and to see where they end up, with implications as to which would be the best sediments to remediate. This might require subscripting certain variables in the code so the calculations could be run in parallel. EPA has started to do some of this, but more would be instructive.

Response 10-EA-1:

In response to the Reviewer's first question, EPA does not expect EFDC to predict how well specific remedial engineering techniques will perform. It is standard practice to use other applications specifically designed for such analyses (e.g., the U.S. Army Corps of Engineers approaches to engineering cap stability). Such an approach was used in both the ½-Mile and 1½-Mile Removals, conducted by GE and EPA, respectively. EPA expects that similar tools will be used when the final design of the selected alternative is performed, not EFDC.

EPA agrees that the second question posed by the Reviewer is one that is addressed by the example scenarios, but is also addressed by the tracer simulations (see Section 4.2 of the FMD).

The Reviewer's comment regarding a "cold start" is unclear, and may represent a misunderstanding of the conditions at the start of these example simulations. Initial conditions assigned in the example simulations were the same as those used in the Phase 1 Calibration (post spin-up), as described in Section 7.1 of the MVR, and therefore do not constitute a "cold start." The only exception was for Example Simulation 2, in which Reach 5A sediment PCB concentrations were set to zero for the purpose of the demonstration.

The deposition of particulate PCBs in Reach 5A is noted in the discussion of Example Simulation 2 (no PCBs from East Branch boundary, no PCBs in Reach 5A sediment, and no bank erosion in Reach 5) in Section 4.1.2 of the FMD. In that simulation, PCBs entering the PSA from the West Branch represent the only source of PCBs to Reach 5A sediment, and contribute to a small increase in PCB concentrations in sediment in Reach 5A.

The tracer simulations presented at the Document Overview Meeting and discussed in Section 4.2 of the FMD provide the ability to analyze the origin and disposition of PCBs in a manner equivalent to the Reviewer's suggestion to evaluate "different colored" PCBs.

Figure 4.2-53 shows areas of net deposition indicating ~ 42% of net deposition is to floodplain, and ~ 10% is to Woods Pond. The remaining ~48% is to the sediment bed upstream of Woods Pond. It would be nice to keep track of the deposited sediments. My guess is that those deposited to the flood plain and to Woods Pond stay put, or are slowly released, while those in the sediment bed move on. Can this be quantified?

Response 10-EA-2:

The Reviewer is correct that PCBs deposited to the river sediment can be released back into the water column. In addition, the model suggests that the floodplain serves as a sink for solids with the exception of erosion from the banks. The tracer simulations provide a quantification of the disposition of PCBs in the river channel and floodplain under a base case and seven different scenarios. Similar quantification could be conducted for solids.

W. Frank Bohlen

The model response for the two examples shown seems reasonable in the sense that most locations experience some decline in PCB concentrations over the upper 6in of the sediment column. I would have expected to see more of a change in Example 2 since virtually all of the major sources of PCB to the system have been shut off. This may point to the importance of the floodplain as a continuing source of PCBs to the system and/or point to the fact that chage can only be expected to occur over depositional depths. These depths, controlled by sediment rates in the area, will seldom exceed 3in except in the vicinity of point bars or similar channel features that are not well modeled because of the coarse grid scales used in the river channel. Since the model is averaging over a significantly greater depth (6in) the changes over 26 years may be difficult to assess. I'd be interested in seeing what a model run of 50 years or more would show.

Response 10-FB-1:

As noted by the Reviewer, the major external source of PCBs to Reaches 5 and 6 was set to zero loadings in the example simulations. The Reviewer comments that contributions from the floodplain may be an important continuing source, and EPA agrees that bank erosion appears to be an important continuing source. An additional internal source of PCBs is sediment resuspension. Both represent continuing sources of PCBs to downstream reaches.

However, it should be clarified that the model results from the vertical segmentation of multiple layers simulated in the bed were averaged to develop the graphical summaries of the example scenarios. The presentation of the results as 6-inch averages (of these layers) was chosen to facilitate the comparison of these results with data (which were typically collected in 6-inch increments).

Longer simulations (e.g., 50 years) will likely be conducted as part of CMS. Under the terms of the Consent Decree this work will be conducted by GE.

Douglas Endicott

The MVR presented 25-year forecasts made with the linked models, in the Rest-of-River and Downstream domains, for 2 (or 3)[1] hypothetical remediation examples. These were intended to illustrate the performance of the models in forecasting the outcome of remediation alternatives in terms of PCB concentrations in sediment, water and especially fish. These forecast predictions ("projections") are really the unique product of water quality models, and are the only defensible means of anticipating the long-term outcome of remedial action and the relative effectiveness of remedial alternatives. We are asked in this question to address the "reasonableness" and "plausibility" of the projections, which are both subjective criteria. Dom DiToro called this sort of evaluation the "laugh test", where we compare the numerical model's projections to those of the models we carry around in our heads. The modeling team is correct in pointing out that it is not possible to determine whether the example projections are "correct" in terms of accuracy (MVR, p 7-3), because there is really nothing to compare them to except one another.

Model projections for the example scenarios were presented a number of ways in the MVR and its supplements. These included the predicted time series of PCB concentrations in water, at the

---

[1] I am somewhat confused by what is referred to as a "base case" projection in a number of the time series figures (e.g., "Time series of water column concentrations for PCBs and TSS", "Time series of sediment PCB concentrations in each spatial bin" and "Time series of plots of FCM exposure concentrations" in *the EPA Response to Questions* document, and "Food chain model results" in the Document Overview presentation of Example Scenarios, which included a graph subtitled "Base case for the example runs - existing boundary loads"). The "base case" scenario is not mentioned or described in Section 7 of the MVR, or anywhere else I could find in the materials provided for our review. This makes it difficult to evaluate the example scenario projections, because I want to compare them to a "base case" (as is intended by the graphs mentioned above) but am not sure what this is. Initially I thought the "base case" was simply the validation prediction, since the hydrograph and boundary conditions are the same. Upon inspection, however, it appears that the initial conditions for PCBs are somewhat different, most noticeably in Woods Pond. A description of the "base case" scenario should be added to Section 7 of the MVR to resolve this ambiguity. Furthermore, EPA should consider whether a scenario should be added or modified which specifies the East Branch PCB boundary condition to be 20 ng/L, the assumed nominal concentration following cleanup.

conventional monitoring locations (Holmes Road, New Lenox Road, …); PCB concentrations in surficial sediment in the mile reaches; and, PCB concentration in largemouth bass in the river reaches 5A thru 6. For EFDC, the projections were also presented as plan maps of the grid displaying PCB concentrations in surficial sediment and soil, and changes in concentration, at the end of the 25 year projections. For the top 6" of the sediment bed, the end-of-projection results are also shown as profiles. Each of these is discussed below.

Response 10-DE-1:

Simulation of the base case used the model inputs and the PCB initial conditions in sediment and floodplain from the Phase 1 Calibration (based on data collected primarily in 1998 and 1999), and hydrograph used for the validation simulation (as described in Section 7 of the MVR).

The differences between the validation simulations and base-case initial conditions noted by the Reviewer reflect the use, for validation, of 1980 data to develop initial conditions for PCBs in Woods Pond and estimates of initial conditions in other reaches, based on the scaling procedures described in FMD Section 2.3.3 and Appendix B.3, and MVR Section 4.2.

The projected timeseries of water column TSS and PCB concentrations are plotted together in the figure "Time series of water column concentrations for PCBs and TSS" in *the EPA Response to Questions* document. The base case projection is qualitatively similar to the validation predictions, with TSS and PCB concentrations increasing to high levels during flow events, although over time the base case PCB concentration projections decline during low-flow periods. PCB concentrations projected for the example 1 scenario are lower than the base case, and projections for the example 2 scenario are lower than for example 1. In each scenario, there is a similar "event responsiveness" of PCB concentrations to fluctuations in the hydrograph. TSS concentration projections are comparable in each scenario, as would be expected since the TSS boundary condition and grain size distribution in the sediment bed are unchanged from the base case. I would consider the water column projections for each of these scenarios to be reasonable and plausible.

Response 10-DE-2

EPA agrees with the Reviewer's assessment that the concentrations computed in the example scenarios are plausible.

The projected timeseries of surficial (top 6") sediment PCB concentrations in river mile reaches are plotted together in the figure "Time series of sediment PCB concentrations in each spatial bin" in the *EPA Response to Questions* document. The base case projections for surficial sediment PCB concentrations are qualitatively similar to the validation predictions, and suffer similar errors. PCB concentrations are projected to increase in the first half-mile reach (an apparent artifact), and show almost no change after start-up in the other reaches. The rate of change in PCB concentrations in sediment is (again) unreasonably slow. Even 25 years is not long enough to see a discernable change in surficial sediment PCB concentrations, which I would certainly expect to see. Nothing shown in the MVR or at the peer review meetings convinces me that the model is correct in its predictions of the rate of change in sediment PCB concentrations. For example scenario 1, a decline in sediment PCB concentrations is projected in comparison to the base case, which is more encouraging. The difference is most dramatic in the first half-mile reach (RM 135.13-134.89), and is also apparent in reaches RM129.19-128.88, 128.88-128.69, and 126.99-126.47. In other reaches, especially in Woods Pond and its headwater, there is no difference in PCB concentrations between example 1 and the base case. For example scenario 2, negligible (1 mg/Kg or less; I can't tell the actual value due to the axis scaling) sediment PCBs are projected for reach 5A due to remediation. In the other reaches, sediment PCB concentrations are projected to be nearly the same or identical to example 1 projections. Although I don't believe that the model's projections of rates of change in sediment PCB concentrations are reasonable, at least the relative differences in projected sediment PCB concentrations between the scenarios are plausible.

Response 10-DE-3:

The Reviewer is correct in noting that the base case results are similar to the validation results for reasons described in the response above. The two simulations use the same inflow hydrographs and boundary conditions, and differ only in the sediment PCB initial conditions.

The Reviewer's comment regarding the artifact of the bed load boundary condition on sediment PCB concentrations in the most upstream spatial bin is also correct. This has subsequently been addressed through a change in the f_{OC} associated with bed load.

It should be noted that the opinion provided by the Reviewer that "The rate of change in PCB concentrations in sediment is (again) unreasonably slow" is not consistent with the results of the statistical analysis of sediment PCB data described in Appendix A.1 of the FMD, the RFI, and other materials provided to the Reviewer by EPA. Appendix A.1 of the FMD provides further statistical analyses of temporal and spatial trends in sediment data.

The Reviewer's evaluation of model results, when placed in context of the determination of lack of statistically significant trend in PCB concentrations in sediment, led to the erroneous conclusion that the simulated rate of sediment PCB concentration decline is too slow. The simulated slow rate of decline in concentrations is well within the uncertainty envelope of the data and totally consistent with the statistical analyses.

EPA agrees with the Reviewer's comment that the relative differences in projected sediment PCB concentrations between scenarios are plausible.

The grid map diagrams (Figures 7-1 and 7-3) show the cumulative change of PCB concentrations in the top 6" of sediment and floodplain soil over the 25 year projection, for scenarios 1 and 2[2]. It appears that the sediments in reach 5A, which are cleaned up in example scenario 2, recontaminate to 1 or 2 mg/Kg over 25 years, due to continuing transport of PCBs from the West Branch boundary condition. It would be useful to see a similar grid map diagram in which the difference in end-of-projection PCB concentrations was plotted.

Response 10-DE-4:

The results of the example simulations were presented in various formats. The Reviewer suggests an additional way of viewing the information from the example scenarios that is potentially useful. This may be considered in displaying model results in the CMS.

The surficial sediment PCB concentration profiles offer a useful alternative view of the cumulative change of PCB concentrations over the 25 year projection. Projections for example 1 show the now-familiar redistribution of sediment PCBs in the high-energy reach 5A, but also significant changes (approaching ± 100% of the IC) in reaches 5B and the upper end of 5C. At the lower end of reach 5C and reach 6, sediment PCB concentrations decline by 20 to 80%. This seemingly contradicts what I interpreted from the "Time series of sediment PCB concentrations in each spatial bin" figures. Could it be that the combination of reach averaging, displayed on a log-axis plotting, makes it difficult or impossible to properly interpret the projection results?

Response 10-DE-5:

The Reviewer comments that changes in sediment concentrations by 20 to 80% in Reaches 5C and 6 in individual grid cells "… seemingly contradicts what I interpreted from the 'Time series of sediment PCB concentrations in each spatial bin' figures."

EPA notes that the changes in concentration over the course of the simulations used in this exercise (base case or example simulations) are consistent with the changes simulated over the course of the validation simulation (MVR Figure 6.2-51). The Reviewer is correct in noting that spatial averaging precludes the ability to view cell-to-cell variability in concentrations, which are displayed on the plan-view figures. Log-scales were used to display the results for the spatial bins, allowing the comparison of all spatial bins on a consistent scale. However, use of the log-scale minimizes the visual difference among the results of the base case and two example simulations. As noted previously, each graphical display has advantages and disadvantages; therefore, the results of the example simulations were presented in a variety of formats.

---

[2] I suppose the comparable graphic for the base case would be the validation figure (Figure 6.2-55) although I am forced to speculate here.

At the Peer Review meeting last month, there were questions and discussion about whether it was reasonable that the EFDC model projected a significant reduction in PCB concentrations in the water column, but little or no change in sediment concentrations. In fact, such behavior is expected in water quality models incorporating both water column and sediment sedimentation. In such models, the concentrations of particle-associated constituents change much more rapidly in water than they do in sediment, whenever there is a change in the constituent loadings or boundary conditions. Since PCB boundary conditions and loadings (actually fluxes from reach 5A sediments) are being manipulated in the example scenarios, the differences in the projections of PCB concentrations in the water column versus sediments are entirely expected and reasonable.

Response 10-DE-6:

EPA agrees with the Reviewer’s comment that differences in the water column and sediment responses are reasonable.

Example scenario projections made by the FCM were presented at the Document Overview meeting (13 - Example Scenarios.pdf) as timeseries of PCB concentrations in largemouth bass. In reach 5A, bass in the base scenario are forecast to have a body burden of 50 mg/kg at the end of the 25 year projection; for example 1, the comparable forecast is 30 mg/kg, and 0 concentration for example 2. In reach 5B, bass in the base scenario are forecast to have the highest PCB body burden at the end of the 25 year projection: 70 mg/kg; for example 1, the comparable forecast is 40 mg/kg, and less than 20 mg/kg for example 2. Further downstream in the Rest-of-River, the remediation examples are projected to considerably less effective in reducing PCB concentrations in comparison to the base case. In reach 5C, bass in the base scenario are forecast to have a body burden of 30 mg/kg after 25 years; for example 1, the comparable forecast is 20 mg/kg, while for example 2 the forecast is 10 mg/kg. At the end of the 25 year projections in reach 6 (Woods Pond), bass in the base scenario are forecast to have a body burden of 50 mg/kg; for example 1, the comparable forecast is 40 mg/kg, while for example 2 the forecast is 30 mg/kg. To me, this progression in the declining effectiveness of remediation as one moves downstream appears reasonable in the context of PCB water and sediment exposure projections for these example scenarios. In “no data” reach 5D, the least reduction in bass PCB concentrations due to remediation is projected. I suppose this may be because this reach is relatively isolated from the rest of the river.

Response 10-DE-7:

EPA agrees with the Reviewer’s comment that the FCM results from the example simulations are reasonable.

Marcelo Garcia

Model projections seem reasonable but I would have liked to see more potential scenarios, particularly for extreme conditions such as low flow summer-like conditions and floods.

Response 10-MG-1:

Both example scenarios were simulated using the 26-year validation hydrograph. This hydrograph includes extreme conditions such as low-flow summer-like conditions and floods.

Frank Gobas

The patterns in the data indicate a slow temporal response of the PCB concentrations in the River. The example scenarios also indicate a small decline of somewhere between 0 to 5% (it is hard to spatially average the concentrations depicted without additional information) over 26 years in response to the assumed loading reductions. I think that this is consistent with the observations.

The data illustrate large small-scale spatial variability in PCB concentrations in sediments. The model appears to capture this to some degree as changes in concentrations relative to the control vary between grid cells within many transects.

I have trouble understanding why PCB concentrations in so many cells of the lower part of the River (Figure 7-1b and 7-3b) increase in concentrations as a result of the loading reductions. This is not what I would expect to happen intuitively given the long history of the PCB contamination problem and the slow response time of PCB concentrations in the river. I would expect concentrations to go down throughout the river, but at higher rates at some locations and lower rates at others.

Comparing scenarios 1 and 2, one would expect that elimination of additional PCB loads in scenario 2 would produce a greater change in PCB concentration over time in scenario 2 than 1. Perhaps, this is the case. It is hard to see from the graphs. However, even if this is not the case, it is possible that concentrations decline over the 26 years are comparable for scenarios 1 and 2. Without more information, it is hard to be more definite.

Based on the current information presented, there is no basis for concluding that the patterns provided in the example scenarios are not plausible given the patterns observed in the data. But more data is needed to support a more positive and definitive conclusion. The example scenarios in the validation report provide little information about the functioning of the model. To address the charge question properly, it is important that the model outcomes of the example scenarios are further analyzed. In particular, it is important to average model outcomes over space and time such that the model predictions can be compared to available data.

Recommendation

I recommend that the behavior of the combined model is explored in greater detail than is presented in the validation report. I recommend that the model outcomes in the model scenarios are aggregated to depict the overall response of the PCB concentrations in the River. This will provide the opportunity to better compare model projections to available data sets and judge whether the model projections are plausible.

Response 10-FG-1:

EPA agrees with the Reviewer's comment that the small decline shown in the example simulations are consistent with the data. The Reviewer commented that small increases in PCB concentrations in downstream grid cells were unexpected. However, EPA believes that these results are expected because similar increases are simulated in the base case. These small increases are computed in floodplain cells that start with very low initial concentrations of PCBs.

Comparing Examples 1 and 2, there is minimal difference in patterns of PCB concentration changes in main channel sediment. Minor increases in PCB concentrations (relative to the initial PCB conditions of zero) occur in Reach 5A. In these simulations, these small increases are caused by PCB inputs from the West Branch.

In response to recommendations made by Reviewers at the MVR Document Overview Meeting, model results were aggregated in the additional presentations of results from the example scenarios included in Attachment 5 to the document, EPA Response to Questions from Model Validation Document Overview Meeting. These included figures showing water column, sediment, and food-chain exposure concentrations, presented in formats consistent with those used to present data in the MVR.

Wilbert Lick

Because the basic processes are inaccurately described and the correct processes may possibly differ by factors of two to ten from those in the present model, the projections of the present model have large potential errors.

Response 10-WL-1:

See General Response above and also the responses to General Topic 1. It must be recognized that the Reviewer conclusions are based on erroneous assumptions regarding the scale of processes such as sediment-water mass transfer flux of PCBs.

E. John List

With respect to the PCB modeling the results are so tenuous (see Figure 6.2-49) that it is not possible to draw any really solid conclusion with respect to bias. However, an important point is demonstrated by Figure 6.2-50. This figure indicates that the sediment concentrations of PCB in the lower reaches of the project study area have declined somewhat since 1990 (i.e., since remediation activity started). Furthermore, there appears to be a similar decline in the water column concentrations of PCB, as is somewhat evident in the lower panel of Figures 6.2-44. (It is noted that there is no statistical measure of the significance of the decline in either the sediment or the water column, but the data are very suggestive). However, the modeling does not give any indication that any similar rate of decline has occurred (see Figure 6.2-50). In fact the comparison between the data and the model in this figure is strongly suggestive that the model does have a bias.

Response 10-JL-1:

The Reviewer states that the data shown on Figures 6.2-49 and 6.2-50 are suggestive of statistically significant declines in sediment PCB concentrations. The data are presented as the mean ± 2 standard errors, which approximates the 95% confidence interval of the mean. At that time, the wide range in the confidence interval of the mean and the small number of samples in earlier years were identified by the modeling team as factors that complicated characterization of trends in the data and comparisons with trends computed by the model.

In response to the misconception of the Reviewers regarding the existence of a trend in PCB concentrations in sediment, a more rigorous statistical analysis was performed (see Appendix A.1 of the FMD and the Response to General Topic 1). This analysis confirms previous conclusions that there is no statistically significant decline in sediment PCB concentrations. The results from this (and prior) analyses contradicts the basis for the Reviewer's conclusion that these model-data comparisons indicate a bias in the model.

The bias of the model is further reinforced by the results of the hypothetical remedial scenarios plotted in the un-numbered figure on page 56 of the EPA Response to Questions from Model Validation (Document DCN: GE-061406-ADFI), where, despite a reduction in the flux of PCB from the sediments in the upper river reaches, there is no significant change in projected sediment concentration of PCBs in the lower reaches. Thus the model does not match the reduction in sediment concentration that has occurred subsequent to actual post-upstream remediation, and projects no reduction as a result of an hypothetical reduction in the upstream PCB flux, as in the hypothetical examples. This problem needs to be understood and/or fixed before the model is used in real applications.

Response 10-JL-2:

The results from the simulations of the example scenarios show a gradual decline in PCB concentrations in sediment in downstream reaches for the base case and Examples 1 and 2 (which include elimination of upstream sources). The differences among the rates of decline in sediment PCB concentrations in the base case and example simulations are small, which is consistent with the statistical analysis that concludes that conditions in the river are changing slowly.

It must be noted that the Reviewer's statement that impacts of remediation should have an appreciable influence on the model results is incorrect. With the exception of limited activity associated with Building 68, the remediation in the ½-Mile Reach began in October 1999, and remediation in the 1½-Mile Reach began in 2002. While the model simulation used the 26-year validation hydrograph, the initial conditions were derived using data that pre-dated remediation. The Reviewer's evaluation of example scenario simulations is based on the incorrect assumption that the data reflect 9 years of remediation activities.

I am not a biologist, but having said that, to a layperson the broad distributions of the PCB in the biota appear to be reasonably well described, at least at the central tendency level. This may well be due to the fact that partitioning of PCB into biota appears to be closely associated with the spread of concentrations in the environment. In the example scenarios there is no projected effect on the mean PCB concentration in the Woods Pond sediments, which does seem surprising, especially given the results in Figure 6.2-50, where there is a seemingly significant drop in the PCB concentration in the sediment between 1990 and 2000. However, there is a projected rough order of magnitude reduction in the water concentration of PCB in Woods Pond as a result of the hypothetical scenarios. The actual data for Woods Pond referenced above seem to show that there is a reduction in both water column and sediment concentrations coincident with the remediation that has occurred so far. It is not clear why the modeling should necessarily uncouple the water concentration from the sediment concentration in Woods Pond in the hypothetical examples.

Response 10-JL-3:

The Reviewer's statement that "In the example scenarios there is no projected effect on the mean PCB concentration in the Woods Pond sediments…" is incorrect. Small declines in concentration were simulated over the course of the 26-year simulation in the base case and the example scenarios. Differences in the rate of decline among the base case and two example scenarios were also simulated. The differences in the rates of decline are small, as is expected in a system in which PCB concentrations are changing slowly.

It should be noted that the modeling does not decouple the response in the water column and sediment, as suggested by the Reviewer. It is expected that much slower responses will be simulated in sediment compared to the water column in a system with a relatively low sedimentation rate, as is the case in Reaches 5 and 6.

11. GENERAL TOPIC 11: FOOD CHAIN (BIOACCUMULATION) MODEL

11.1 COMMENT SUMMARY

Reviewer Name	Page Number/Line Number
Bohlen	Page 15, lines 13 to 15
Endicott	Page 9, lines 4 to 27
	Page 10, lines 1 to 4
	Page 10, lines 24 to 27
Gobas	Page 4, lines 17 to 18 and page 4, lines 42 to 43
	Page 5, lines 29 to 31
	Page 5, line 44 to page 6, line 2
	Page 6, line 42 to page 7, line 17
	Page 7, lines 38 to 42
List	Page 6, lines 1 to 5

11.2 GENERAL TOPIC SUMMARY

The model framework for the Housatonic River is comprised of three models, a watershed model (HSPF), a hydrodynamic, sediment transport, and PCB fate model (EFDC), and a food chain bioaccumulation model (FCM). This section provides responses to comments regarding FCM.

FCM was calibrated in two ways. The first was to perform a calibration of the model using the data. This calibrated model was then used in a linked mode (receiving exposure concentrations from EFDC and temperature data from HSPF). The results from the linked simulation were evaluated and no additional calibration was necessary.

It should be noted that many of the comments regarding FCM were related to the presentation and/or interpretation of the uncertainty analysis. These comments are discussed under General Topic 5.

Three Reviewers commented that FCM yielded estimates of PCB central tendencies in organisms that were in reasonable agreement with data. Three reviewers made statements regarding the suitability of the model-data comparisons used to evaluate FCM model performance.

11.3 GENERAL TOPIC RESPONSE

EPA agrees with the general conclusion expressed by several Reviewers that model/data comparisons of average PCB concentrations in fish tissue using the results from FCM are in reasonable agreement. In addition, these simulations show little or no systematic bias. The FCM Calibration and Validation (both linked and data-driven) in Reaches 5 and 6 and also the simulations in Reaches 7 and 8 all generated overall model bias statistics that were close to the zero bias value of 1.0. The range of overall model bias values (MB*) for the above modeling components ranged from 0.88 to 1.11, which demonstrates an unbiased model, particularly in light of the uncertainties in measured fish tissue and exposure concentrations. EPA disagrees with some of the Reviewer comments made regarding perceived bias and/or believes that the perceived bias is overstated. In some cases it appeared that Reviewer comments did not reflect all of the information provided (including the EPA Response to Questions from the Document Overview Meeting). These comments are discussed individually below.

11.4 RESPONSE TO SPECIFIC COMMENTS ON GENERAL TOPIC 11

W. Frank Bohlen

Finally, I'd consider eliminating the FCM in favor of a parametric (flow, TSS concentrations ?) approximation of body burden uptake based on the results of the complete model runs.

Response 11-FB-1:

This comment was made in the context of a suggestion to reduce the overall computational burden of the model framework. EPA strongly disagrees with this recommendation for two reasons. First, replacement of the mechanistic food chain bioaccumulation model with a black-box empirical model would be inconsistent with the need to represent the important biological processes (balance of fate processes) that govern PCB uptake in the Housatonic River. In response to recommendations from Reviewers made during previous Peer Reviews, a tiered approach was used to identify, characterize, and parameterize the bioaccumulation and fate processes necessary to achieve the goal of the modeling study. EPA believes that to deviate from this analysis and eliminate these important processes would compromise the scientific integrity of the linked model application. Second, the elimination of FCM would not have an appreciable effect on the computational requirements of linked model runs. The FCM code operates on a level of spatial and temporal resolution that is orders of magnitude coarser than EFDC, requiring only minutes for each simulation.

Douglas Endicott

The food chain model (FCM) predictions should be compared to the mean or other central tendency of the fish species, reach, and age-specific data, because the model is designed to simulate "average" fish according to these categories. This is confirmed by analyses and graphics (e.g., Figure 4.3-8) displaying that the residuals tend to be smaller when the data are averages of 6 or more fish. Unfortunately, many calibration and validation graphics persist in using individual data.

Response 11-DE-1:

EPA agrees with the Reviewer that model-data comparisons should emphasize central tendencies for unique combinations of fish age, species, and river reach. Unfortunately, the sample sizes for such comparisons are sometimes insufficient to conduct meaningful evaluations using this approach. Therefore, while the central tendencies in calibration and validation graphics and statistics were emphasized, individual samples were reported in model-data comparisons where only individual data or small sample sizes were available. This approach was considered preferable to the censoring of data with small sample sizes because it provided a comprehensive analysis of the available data. Figure 4.3-8 of the MVR provides an example of such a combined presentation, allowing the reader to differentiate between comparisons made using either small or large sample sizes.

Nonetheless, there is overall good agreement between predictions and the averaged total PCB concentration data, with most residuals falling within a factor of 2 of the predicted values. EPA has revised Figure 4.3-7 to use the 95% confidence limits (or ± 2 standard errors) to better quantify the measurement precision, and the FCM predictions fall within these limits. FCM model performance is similar in calibration and validation periods.

Response 11-DE-2:

EPA agrees with the Reviewer comments. The "factor of 2" mentioned by the Reviewer was selected a priori as the model performance measure for model Calibration; therefore, the Reviewer's observation supports EPA's conclusion that the model is adequately calibrated. Furthermore, the model performance during Validation was conducted using a model that was not changed from the Calibration, and applied a fully independent data set. It should be noted that the a priori model performance measure for Validation was a factor of 2.5. The similar quality of model performance observed between calibration and validation periods provides additional evidence of model robustness and assurance that the model is adequate for use in achieving the goal of the modeling study.

Figure 3-9 (Comparison of mean measured biota tissue tPCB concentrations to FCM results [simulation using linked models]) provides a summary illustration of the FCM calibration. PCB concentrations predicted for invertebrates are in good agreement with the D-Net sample data, except for infauna in Reach 5A (which are substantially overpredicted).

Response 11-DE-3:

EPA agrees with the Reviewer's observation that the FCM simulations for invertebrates are in good agreement with the D-Net sample data. As discussed in the MVR, the D-net data contained primarily epifaunal invertebrates and therefore cannot be directly compared to the infauna simulations. EPA disagrees that the simulation for Reach 5A infauna can be described as "substantially overpredicted." The reason is that there are no field data specific to infauna to assess this. Therefore, the accuracy of model simulations for infauna is unknown, not biased. EPA acknowledges that the PCB concentrations in Reach 5A organisms exposed via sediment are more uncertain than in other reaches due to the atypical partitioning behavior of PCBs in this reach.

The model fits for bullhead and sunfish are very good; PCB concentration in these fish vary little between river reaches. Model-to-data comparisons are not as good for suckers, cyprinids and bass, although the PCB concentration data for these species is also considerably more variable. I assume that at least some of the variability of PCB concentrations in fish reflects the wide range of PCB exposure and/or failure of the food chain to spatially average these exposures on a reach basis.

Response 11-DE-4:

The model fit for all representative species is well within the model performance measures specified for Calibration and Validation; however, the Reviewer is correct that model fit is stronger for bullhead and sunfish relative to other FCM representative species. As noted in the MCR and MVR, the model-data comparisons for cyprinids must be interpreted with caution because FCM simulates a generalized cyprinid in downstream reaches whereas the data collected were for golden shiners. This species has a different life history than the generalized cyprinid simulated. The data for golden shiners also exhibited lower than average lipid content. These factors contributed to the differences observed in the model-data comparisons.

EPA agrees with the Reviewer's observation that the wide range of PCB exposures within any given reach may explain some of the variability of PCB concentrations in fish. Fish integrate their exposures over portions of a reach. However, many fish are territorial and are exposed to localized portions of a reach that may have higher or lower sediment PCB concentrations relative to the reachwide averages used in the model. For example, EPA evaluated the home range of largemouth bass relative to the spatial distributions of PCBs in sediment. Order of magnitude differences in exposure concentrations are expected for fish occupying non-overlapping territories within a river reach.

If I look for an overall pattern in the residuals on this figure, it would be a tendency for the predicted PCB concentrations to increase moving downstream, which is not reflected by the data. PCBs predicted in Reach 5D for cyprinids, sunfish and bass significantly exceed the measured (as well as predicted) concentrations for these species in other reaches. Since PCB concentrations were not measured in fish from this reach, these predictions are untested. Reach 5D is the only portion of the FCM where predictions based on PCB exposures derived from data meaningfully diverge from predictions based on EFDC exposures.

Response 11-DE-5:

EPA disagrees that there is a systematic tendency for simulated PCB concentrations to increase with distance downstream. The slight increase in concentrations in Reach 6 (Woods Pond) is due to the higher PCB concentrations in Woods Pond surface sediment relative to upstream reaches. Both the model and data indicate relatively flat (slightly declining) trends in fish tissue concentrations across the PSA. The PCB concentrations simulated by FCM demonstrate the same trends (within the uncertainty of the data) as the statistical analyses provided in the RFI and in Appendix A.1 of the FMD.

EPA agrees that the PCB concentrations simulated in Reach 5D (for the linked model) are greater than in adjacent reaches and that these simulations cannot be tested due to the lack of fish tissue data for Reach 5D. The uncertainty in fish tissue and exposure concentrations for Reach 5D is greater than in other reaches.

FCM predictions in the downstream model domain are compared to PCB concentrations measured in fish for Reach 7 and Rising Pond. These predictions look favorable overall, although there may be bias in the predictions for specific species. Results appear comparable in the data-based and linked exposure predictions.

Response 11-DE-6:

EPA agrees that the model predictions for Reaches 7 and 8 are favorable and that the linked and data-based applications of the model yielded similar results. Revised model performance statistics are presented in Sections 3.6.4 and 3.6.5 of the FMD (Table 3.6-11 and Table 3.6-13). These statistics indicate that all species with a sample size >5 exhibited model bias (MB*) measures between 0.72 (perch in data-driven model) and 1.44 (bullhead in linked model). These results indicate only a slight bias. Overprediction of bluntnose minnow PCB concentrations was also observed; however, this was based on only five samples, includes an interspecies extrapolation, and the MB* values were less than 2.0 for both the linked and data-based model. The model bias for all species was within the model performance measure for Model Validation.

It is unclear whether the low-flow PCB bias has led to food chain miscalibration, as suggested by GE, because the food chain model was initially calibrated using data-based exposure concentrations. These would not exhibit the same low-flow bias as exposure concentrations predicted by the model.

Response 11-DE-7:

The Reviewer has correctly identified that the data-based FCM calibration results provide an independent assessment of the reasonableness of bioaccumulation predictions. Because both the data-based and linked model calibrations produced similar acceptable model performance, and because no modifications (i.e., process or parameter changes) were made between these calibrations, there is no evidence of “miscalibration.”

During the Validation Peer Review, GE’s consultant incorrectly implied that calibration of the bioaccumulation model was used to compensate for bias in model linkages. As described in the response to General Topic 1, all models in the framework were calibrated independently as presented in FMD Figure 1-11. Information was passed forward from one model to the next in sequence as presented in FMD Figure 1-11. Outputs from HSPF were used as inputs to EFDC and FCM. Further outputs from EFDC were used as inputs to FCM. Information was never passed backwards from FCM to EFDC or HSPF. No compensation was made in one model for a performance issue in another model. For example, FCM was not recalibrated to achieve acceptable model/data comparisons in fish tissue compensate for the any potential bias in flow, TSS, or PCBs outputs from EFDC or HSPF.

A quantitative bounding calculation was performed to evaluate the impacts of underestimating TSS and PCB concentrations at low flows, which results in offsetting effects on calculated exposure concentrations of PCBs sorbed to particulate organic matter in the water column, thereby minimizing the concern over these features of the boundary estimates. This bounding analysis showed that replacing all water column PCB concentrations less than 0.06 µg/L (scaled up to account for overprediction) with 0.0 µg/L PCBs had little impact on the simulation of PCB concentrations in fish tissue (less than 5%, averaged across species, and less than 8% in game fish).

Frank Gobas

The spatial resolution of the environmental fate and food-web model is unnecessarily complex….Also, the time-step requirement for the hydrodynamic model is too onerous and unnecessary for the environmental fate and food-chain models.

Response 11-FG-1:

The Reviewer is incorrect in his assumption that the spatial resolution of the environmental fate (EFDC) and food-web (FCM) models are the same; in fact they are very different. The model grid for EFDC was optimized to provide output at the scales of individual grid cells (20 to 70m), spatial bins (~1 to 2 km), or for a subreach/reach. FCM provides output at the subreach and reach scale. These spatial scales of model output bound the spatial scales of interest to achieve the goal of the modeling study. The Reviewer is also incorrect in his assumption that the time-step requirement for the hydrodynamic model has any relationship to that used in FCM. The temporal resolution (daily time step) of the FCM exposure linkages are much larger than those used in the hydrodynamic and environmental fate and transport model (EFDC).

Due to the heavy reliance on calibration during model development, the relative PCB concentration data in water, sediments and biota that are calculated by the model are overall consistent with the model predictions of the mean concentrations.

Response 11-FG-2:

Although EPA agrees that concentrations of PCBs in biota are consistent with model predictions of mean concentrations, the Reviewer is incorrect in attributing such agreement to “heavy reliance on calibration” in FCM. In fact, as specifically stated in the MFD and MCR, the overall modeling approach for FCM was to minimize the level of parameter calibration from the initial best estimates. Furthermore, because FCM remained essentially unchanged across Phase 1 Calibration, Phase 2 Calibration, and Validation, this provides confidence that model performance is not attributable to overcalibration.

The FCM model can be expected to properly estimate the relevant contributions of water and sediment concentrations as sources of PCB bioaccumulation in benthos and fish species from PCB concentrations in water and sediments delivered by the EFDC model.

Response 11-FG-3:

EPA agrees with the Reviewer’s comment.

Secondly, the combined or linked model (Table 6-4.7) shows a systematic underprediction of the mean PCB concentrations as high as a factor of about 2 for some of the species. Overall, including all species, the underprediction of PCB concentrations in biota is about a factor of 1/0.60 or 1.67. This systematic bias does not appear to be due to the bioaccumulation model itself. Judging from Table 6-4.6 (Model Validation Report), the bioaccumulation model itself appears to have little or no systematic bias.

Response 11-FG-4:

EPA agrees that the results of the linked model for Reaches 7 and 8 summarized in Table 6.4-7 of the MVR exhibit a systematic underprediction that was attributable to the exposure linkages from EFDC. However, the downstream model calibration was subsequently reevaluated and revised output was generated. These results are presented in Section 3.6 of the FMD. Based on changes made to the model for Reaches 7 and 8, the overall model bias

> changed from MB*=0.60 to MB*=1.08; therefore, the systematic under-prediction is no longer present. EPA agrees that the data-based FCM model indicates little or no systematic bias.

Judging from Figures 6.3.3 to 6.3.8, the uncertainty in the characterization of the model bias is large. The authors could have calculated the standard deviation of the mean, but did not do this. If they would have done this, they would have found that the standard deviations of the model bias for any of the PCB concentration in fish data sets are quite substantial due to the large variability in the observed PCB concentrations in the biota. This means that while the mean model bias for the PCB concentration predictions in fish is relatively low, the uncertainty of the mean model bias is high. I think that it is important in any model to be upfront about the ability of the model to make predictions of reality. The reporting of the model bias without its uncertainty is misleading in my view. The reality is that PCB concentrations in biota vary substantially and we do not really understand why this is. So, we should not pretend that we can predict PCB concentrations in fish with the accuracy that the mean model bias measures suggest. I therefore suggest that the authors provide a full reporting of the model bias of the PCB concentrations (i.e. report uncertainty in the mean model bias) using the linked model and interpret the findings in terms of model uncertainty when the model is applied.

> **Response 11-FG-5:**
>
> In response to the Reviewer’s comment, measures of the variability around the mean model bias have been included in Section 3.8 of the FMD. The 95% confidence limits about the mean model bias were calculated for calibration and validation applications. Further discussion of comments related to FCM uncertainty is provided in the response to General Topic 5.

I recommend that the authors calculate the standard deviations of the mean model bias of the linked model and make appropriate corrections to account for the systematic underprediction of the mean PCB concentrations in the biota by the model. The systematic underprediction of the mean PCB concentrations in fish by the model needs to be either corrected or recognized when the model is applied.

> **Response 11-FG-6:**
>
> See responses above. There is no longer any systematic underprediction of mean PCB concentrations in fish by the model. Therefore, there is no need for “correction” or recalibration of FCM.

E. John List

I am not a biologist, but having said that, to a layperson the broad distributions of the PCB in the biota appear to be reasonably well described, at least the central tendency level. This may well be due to the fact that partitioning of PCB into biota appears to be closely associated with the spread of concentrations in the environment.

Response 11-JL-1:

EPA agrees with the Reviewer’s comments.

12. GENERAL TOPIC 12: USE OF THE MODEL TO EVALUATE REMEDIAL ALTERNATIVES

12.1 COMMENT SUMMARY

Reviewer Name	Page Number/Line Number
Adams	Page 11, line 43 to page 12, line 3
	Page 14, lines 40 to 45
Bohlen	Page 2, lines 16 to 35
Endicott	Page 22, lines 5 to 30
Garcia	Page 1, line 39 to page 2, line 18
	Page 3, lines 15 to 26
Gobas	Page 9, line 29 to page 10, line 25
	Page 13, lines 33 to 39
List	Page 6, line 22

12.2 GENERAL TOPIC SUMMARY

One of the purposes of the modeling study required by the Consent Decree is to provide a tool that is adequate to assess relative performance among proposed remedial alternatives and against baseline conditions. Under the terms of the Consent Decree, after completion of the Model Validation Report Peer Review, EPA is to provide the model and all inputs and outputs from Model Validation to GE for its use in evaluating remedial alternatives in the Corrective Measures Study (CMS).

Reviewers expressed comments on the use of the model for evaluating remedial alternatives. Two common themes were noted, the first was that four Reviewers believed that the model was either incomplete or not good enough to be used to assess remedial alternatives. There was also the sense that the Reviewers expected EPA to provide further information with regard to model performance for specific remedial alternatives. The Reviewers' specific comments and responses are provided in Section 12.4.

12.3 GENERAL TOPIC RESPONSE

EPA believes that the model framework developed for the Housatonic River is adequate for projecting baseline conditions and evaluating the relative performance of potential remedial alternatives. The justification supporting this conclusion is summarized below, and is further described in other responses to comments and in the FMD.

- Model-data comparisons demonstrate that the models accurately simulate flow and concentrations of solids and PCBs in sediment and fish tissue. The Reviewers' comments that the model did not predict an "apparent" decline in PCB concentrations in surface sediment are incorrect because the "apparent" decline is not statistically significant (see General Topic 1). Instead, the model accurately simulates the trend in PCB data, that is, the model accurately predicts that there is little to no decline in PCB concentrations. The simulated concentrations are within the uncertainty envelope of the data.
- The models were parameterized using site-specific data to the extent possible. Model calibration was limited only to those parameters that could not be completely defined using site data. Model Validation and application to Reaches 7 and 8 provide an independent test of model performance and robustness, because the model parameterization was not changed between calibration and these simulations.
- The models accurately predict average PCB concentrations in water, sediment, floodplain soil, and aquatic biota. Model performance also demonstrates that the model framework appropriately characterizes the balance of fate processes affecting PCB concentrations and exposures over time.
- In virtually all cases, the results of the three models in the framework were well within the model performance measures established a priori. The models accurately simulate the central tendencies of the data, and show negligible indications of bias. Detailed sensitivity analyses indicate that the models are responding appropriately to the processes that control PCB transport and fate.
- The most significant of the relevant pathways describing PCB fate in EFDC are upstream boundary loads, deposition onto the floodplain, bank erosion, and PCB loss by erosion of sediment in Reach 5A as described by the process-based flux summaries, example scenarios, and tracer analyses. These analyses confirm the balance of processes simulated in the model and provide additional verification of model performance.

Therefore, EPA believes that the model framework is a tool that can be used as one line of evidence, with other information including data analyses and the risk assessments, to achieve the goal of the modeling study.

As noted above, it was explicitly specified in the Consent Decree that it was not EPA's responsibility to use the models in the three Peer-Reviewed phases of model development specified in the Decree, the Modeling Framework Design, Model Calibration, and Model Validation, to either project baseline conditions or to evaluate remedial alternatives. Therefore, such analyses were not performed.

12.4 RESPONSE TO SPECIFIC COMMENTS ON GENERAL TOPIC 12

Eric Adams

As I expressed at the DRM, I am concerned with the absence of validation data appropriate to remedial measures. Dredging would result in horizontal gradients in sediment concentration over 10s to 100s of meters, while capping would result in vertical gradients in sediment concentrations over a few cm. The current sediment concentration distributions do not show variability over these scales, so it is difficult to tell whether or not the model will be able to successfully predict the effects of such remedial measures.

Response 12-EA-1:

It is not the state of the practice to have collected remediation data to validate the model before making a decision that the model provides an adequate tool for use in evaluating the relative response of the river system to potential remedial alternatives versus baseline conditions.

Considering the difficulty of the modeling task, the current state-of-the-art, the amount of effort that has already gone into the modeling, and the difficulty in arriving at a consensus (when our review panel is not allowed to), I would say that ***the current work is "acceptable"***. However, in view of the many reservations that many of us have, the model users (GE) will have to exercise considerable professional judgment in their use of the model, and EPA should grant them this judgment.

Response 12-EA-2:

EPA agrees that the model is acceptable for use in evaluating remedial alternatives. It is GE's responsibility to submit a Corrective Measures Study (CMS) Proposal for EPA review and approval, and then to subsequently prepare the CMS, again for EPA review and approval. EPA expects to work with GE to discuss questions as they arise in the preparation of these documents, including those related to application of the model in the CMS.

W. Frank Bohlen

This has been and remains an ambitious project. In hindsight it may have been overly ambitious to expect a single model formulation to efficiently and accurately predict future PCB concentrations throughout a morphologically complex region over extended periods of time given what was known about historical distributions and the range of processes governing transport and fate. Some of this seems to have been recognized by EPA in their recent refinement of the modeling goals placing primary emphasis on the need for the model to be able to establish the relative performance of selected remedial alternatives rather than on its ability to yield certain prediction of absolute PCB concentrations (EPA, 2006). While this refinement is advisable and will be taken into consideration in model evaluations it must be remembered that it does not relieve the modeler's responsibility to develop an efficient, stable and quantitatively accurate model. The evaluations to be conducted during the upcoming Corrective Measures Study (CMS) by GE will be considering the benefits of remedial alternatives over extended periods of time, often 40 years or more, i.e. times well in excess of the validation period. The utility of the relative comparisons over these extended periods will ultimately depend on the degree to which the model provides accurate simulation of all of the governing physical chemical and biological factors affecting transport and fate and the adequacy of the computational numerical schemes. Fundamentally, these are the same factors to be considered if the model was to be used for absolute predictions. The sufficiency to evaluate these characteristics depends in large part on our understanding of the PCB transport system structure and dynamics within the Housatonic River.

Response 12-FB-1:

The modeling team agrees with the Reviewer's comments about the complexity of the Housatonic River system and the challenges that this presents for modeling. Model simulations that will be conducted in the CMS will benefit from recent advances in computer hardware since the time of the MVR Peer Review, which have resulted in 33% reductions in run times (see Response to General Topic 4 and Appendix B.1 to the FMD). The modeling team believes, as stated in the General Response, that the model accurately simulates the fate, transfer and bioaccumulation of PCBs and therefore is sufficient to address the goal of the modeling study and is adequate for use in the CMS study.

Douglas Endicott

I interpret this to be more than one question, so I will give two answers. If the goal is to forecast PCB concentrations in the Housatonic River in the absence of further remediation, then the current models may be adequate (i.e., "good enough"). In their current form, the EPA models (and probably many much simpler models) can provide meaningful answers to many questions about the future of PCBs the Housatonic River. The model calibration and validation show the relative magnitude of PCB inputs to system, and the resulting predictions and diagnostics are instructive in terms of evaluating remediation targets. For example, eroding banks and sediment deposits in reach 5A are clearly targets for remediation. The 25-year validation and example scenario projections also address the need and possible effectiveness of remediation in the Housatonic River. Following the remediation of the East Branch reaches, upstream PCB concentrations of 20 ng/L and 0.4 mg/kg expected; these concentrations alone may support gamefish concentrations of 1 ppm or more in the PSA. In the absence of further remediation, PCB concentrations will decline very slowly from the current levels. Although water-column PCB concentrations will decline downstream of remediation conducted at the reach scale, surficial sediment PCB concentrations will be relatively unaffected in downstream reaches. The predicted response of PCB concentrations in fish to remediation is expected to be somewhat similar to the sediment. To substantially reduce (e.g., >50%) PCB concentrations in Reach 5D and 6 gamefish over 25 years appears to require remediation well beyond the spatial extent of hypothetical example 2. None of these conclusions require a highly accurate model.

However, if the goal is to forecast PCB concentrations to evaluate the effectiveness of further remediation, and to distinguish between the effectiveness of various remedial alternatives, then I believe that the current models are not adequate. Although they are good, they are not as good as they could be. I assume that GE will not willingly undertake additional remediation activity in the Housatonic River, unless EPA can convincingly demonstrate that such remediation is necessary and that the remedy will be effective. At their current stage of development, the EPA models are not ready to make such a demonstration. I am arguing that further model development and testing is warranted, and that EPA should be motivated to do so because the financial stakes are high. Otherwise, GE will likely exploit weaknesses in the models to argue against the need for further remediation.

Response 12-DE-1:

EPA believes that the modeling framework is adequate both to describe baseline conditions and to assess relative effectiveness of potential remedial alternatives, as stated in the General Response above. Beyond that, the Charge to the Panel did not include discussion of strategic considerations or motivations of the Consent Decree parties.

Marcelo H. Garcia

It is not enough to state that a given process is accounted for when the representation of the process is deficient and its implementation in the model is done with algorithms and assumptions that are not based on the physics of the process and cannot be supported by field observations. For instance, streambank erosion has been recognized as a very important process since large amounts of PCB-laden sediments can enter the river during medium to high flows. However, fluvial erosion and mass failure of the stream banks are currently modeled simply as functions of flow discharge, a rather crude approximation, which severely limits the capabilities of the model to assess the effect of this process.

How can one determine what regions of the banks are more prone to erosion and, therefore, need protection if the model does not compute the distribution of flow velocity and bed shear stress along the banks? Could native vegetation or other bioengineering techniques be used to protect the banks of the Housatonic River against erosion or more hard-core solutions such as rip-rap are needed? Is it possible to use a combination of both? As it stands now, the model can not be used to compare the merits of such remedial alternatives, which is one of the goals of the modeling study as stated above.

Response 12-MG-1

EPA does not expect that EFDC will be used to evaluate specific localized areas of bank erosion or the efficacy of particular remedial engineering approaches such as use of native vegetation or hard engineering. There are other more applicable and widely used techniques available. Application of these techniques will inform specific engineering considerations, if appropriate, during the final design of the selected remedial alternative. This was the approach used in both the ½-Mile and the 1 ½-Mile Reach Removal Actions.

Due to the nature of the problem and the scale of the river system under consideration, it is unlikely that a full validation of the model will be possible in the near future. From a practical point of view, however, the main question in my mind would be: is the model good enough to capture most of the process responsible for the transport and fate of PCBs so that it can be used to assess the merits and drawbacks of different remediation strategies in the Housatonic River?

Right now my answer would be no but the effort by EPA and its consultants has provided a good foundation towards the goal of having a useful tool that can be used to help the Housatonic and other rivers experiencing similar problems.

Response 12-MG-2:

See General Response above.

Frank Gobas

1. While the report presents several efforts to calculate uncertainty, the model uncertainties are not fully acknowledged. In essence, the report presumes that the model's ability to calculate mean concentrations is sufficient to address the goals of the study. I think that this is a flaw in the study design because the goal of the model is to compare PCB concentrations resulting from different remedial scenarios. Such an application involves the comparison of mean concentrations. However, comparing mean concentrations alone is insufficient to determine the significance of the differences in mean concentrations. The calculation of the statistical significance of a difference in the means is required. The latter is a well established practice in scientific and engineering studies. I do not think that a convincing rationale is presented for why this practice is not applicable in this study. The authors argued that ecological receptors (including fish), due to their continuous movement, tend to be exposed to a large variation in concentrations, which get "averaged out" to produce an internal concentration in the fish that corresponds to the average or mean exposure concentration rather than the variation in concentrations (l. 10-14, p.15 Responsiveness document). Figures 6-2-35, 6-2-56, 6-3-3 to 6-3-8 (Model Validation Report), and similar figures in the calibration documents demonstrate a one-to

two order variability in predicted concentrations of PCBs, which does not differ substantially between fish and sediments (i.e. the variability in fish concentrations does not appear to be any less than that in suspended solids). This indicates that the original assumption that PCB concentrations in fish may be less variable than the sediment concentrations may not hold.

The model's capability to estimate only mean concentrations, makes it difficult to apply the model to some objectives. For example, when the model is applied to address model objective #6, it can only calculate at what point in time the mean concentration falls below a target level that no longer poses a human health or ecological risk. This means that roughly (depending on the frequency distribution of the concentrations) half the concentrations are still above the target level. Risk assessment calculations typically depend on the distribution of the concentration and set limits based on a percentage of individuals exceeding a particular concentration. Hence, a frequency distribution of the concentrations is essential. Perhaps it is assumed that the risk assessment calculations can deal with the distribution of the concentration. However, in the application of the model, it is the model that has to generate the distribution of concentrations (since there are no data for the future) and at this point, none of the model components can do this. A comparable argument can be made for the application of the model to model objective #1, i.e. quantify future spatial and temporal distribution of PCBs (both dissolved and particulate forms) within the water column and bed sediment. The mean concentrations that will be produced by the model do not provide information on the statistical distribution of the predicted concentrations. Hence, as the model stands, it is impossible to determine whether any calculated difference in concentration (e.g. as a result of a remediation strategy) is significant and can be treated as a difference in effectiveness among remediation strategies.

Response 12-FG-1:

EPA disagrees with the Reviewer's comment that the uncertainties of the models have not been fully acknowledged. This was done throughout the MCR and formally in the MVR. EPA strongly believes that the model must only be expected, as an approximation of reality, to reasonably simulate the average concentrations measured in the system being studied. Such models can be appropriate in the sense that they are useful, practical, and scientifically defensible representations of the environment that incorporate a site-specific understanding of system behavior. To this end, the model evaluation is centered on determining "whether the simulation model is an acceptable representation of the real system, given the purpose of the simulation model" (Kleijnen, 1999).

EPA also disagrees with the Reviewer's assumption regarding the interaction between the risk assessment and the model results. The interim media protection goals considering the risk assessments are typically expressed in terms of exposure to the 95% of the UCL of the mean concentration.

The model as it currently stands is incomplete. A lot of excellent work has been done but there are some major gaps that need to be addressed before the model is ready for application. This may sound disappointing to some, especially for those living in the immediate vicinity of the River. However, the remediation options that can be expected to be considered have very large and long lasting impacts on the River and its ecology. Therefore, caution should be exercised and there should be confidence in the outcome of remediation efforts before such remediation takes place.

Response 12-FG-2:

EPA does not believe that there are any gaps in information that prevent the model framework from serving as an adequate tool for use in evaluating baseline conditions and the relative effectiveness of potential remedial alternatives.

E. John List

I do not think so. For the reasons stated above. (above refers to comment 10-JL-2)

Response 12-JL-1:

See General Response above.

13. GENERAL TOPIC 13: OTHER COMMENTS

13.1 COMMENT SUMMARY

Reviewer Name	Page Number/Line Number
Adams	Page 6, lines 7 to 14
	Page 14, line 45 to page 15, line 39
Bohlen	Page 2, line 36 to page 4, line 17
	Page 4, line 34 to page 5, line 19
	Page 9, lines 24 to 25
	Page 10, lines 3 to 14
	Page 14, line 32 to page 15, line 15
	Page 15, lines 16 to 23
Endicott	Page 1, lines 27 to 34
	Page 22, line 31 to page 24, line 4
Garcia	Page 6, lines 37 to 40
Gobas	Page 4, lines 1 to 5
	Page 5, line 21 to page 6, line 2
	Page 8, lines 34 to 37
	Page 14, lines 21 to 32
Lick	Page 2, lines 5 to 7
	Page 10, lines 1 to 5
	Page 13, lines 28 to 29
List	Page 1, lines 5 to 28
	Page 4, lines 9 to 17
	Page 5, lines 5 to 25
	Page 6, lines 31 to 44

13.2 GENERAL TOPIC SUMMARY

This General Topic is intended to provide responses to comments from Reviewers that did not relate to one or more of the preceding General Topics, but were not deemed of sufficient magnitude to justify the creation of an additional General Topic area, and/or were not repeated by other Reviewers. Thus, it is in effect a series of specific responses that are not necessarily related to each other but are organized into one area of the Responsiveness Summary to allow readers to quickly review additional comments and their responses.

13.3 GENERAL TOPIC RESPONSE

Because these comments are not necessarily related to each other, no general response is provided.

13.4 RESPONSE TO SPECIFIC COMMENTS ON GENERAL TOPIC 13

Eric Adams

I believe that several processes could be better represented. But I recognize that several of these processes have been calibrated, so that ***if one parameter were changed, others would have to be changed*** as well. If there is time, I would like to see the model recalibrated using more appropriate values, within physical/chemical/biological constraints. (I do not think an independent validation is necessary.) In the absence of re-calibration, the model users must exercise considerable judgment in the interpretation of model output.

Response 13-EA-1

In response to comments from the Reviewers, some parameter changes have been made. The revisions are noted in the Final Model Documentation Report (FMD) and elsewhere in this Responsiveness Summary. Revised model results and model-data comparisons, reflecting these changes, are provided in the FMD.

And depending on how much time is available, ***I would recommend some modest additional effort***, which I organize into three categories: easy, intermediate and difficult. I believe that GE should be involved with this as much as possible (rather than simply be given a black box), because it is their model judgment that will be relied upon, and this judgment can only come with experience using the model.

Easy

These are things that should be easy to accomplish within a framework of a few weeks to a couple of months. They are probably also the most important.

1. Develop a strategy to improve model efficiency so that remediation scenarios can be evaluated efficiently. I suspect that the best approach here is to make use of synthetic hydrological sequences that are repeated.

2. Perform additional uncertainty analysis that looks at the uncertainty of future remediation scenarios (as opposed to simply past natural attenuation), both in terms of absolute concentrations and in terms of PCB fluxes from different sources.

3. Improve the “upstream” and “downstream” models through further calibration (GE indicates that additional data are available for this purpose).

Response 13-EA-2

1. The application of the model in the CMS will be discussed in the Model Working Group and subsequently proposed by GE and reviewed by EPA. Techniques for improving efficiency, such as the one described by the Reviewer, may be considered.

2. The uncertainty approach presented in the Model Validation Report was intended to be demonstrative rather than prescriptive, and EPA recognizes that other approaches to evaluating model uncertainty may also be valid. The comments from Reviewers on this topic ranged from quite favorable to quite critical, so clearly, no consensus was developed. In subsequent discussions with the Model Working Group following the Peer Review of the Model Validation Report, it was agreed that the uncertainty of the model results in evaluating the relative performance of remedial alternatives would be addressed using the best professional judgment of the Model Working Group. The uncertainty associated with simulations performed for potential remedial alternatives in the CMS will be discussed.

3. The EPA modeling team has further evaluated the necessity for the application of the upstream model and discussed it within the Model Working Group. The decision has been made not to pursue the use of the upstream model (see responses to General Topic 8 for additional details).

In response to the Reviewers’ comments, additional work on the downstream model has been performed. The model was more extensively calibrated, and the final model validation run is presented in Section 3.5 of the FMD (see responses to General Topic 9 for additional details).

Intermediate

The time frame here might be several months to half-a-year.

4. Develop a more highly resolved grid within at least a portion of the channel and use the grid in a short term simulation to understand the sensitivity of model predictions (again, actual concentrations as well as fluxes) to grid resolution. Information parameterized from this sensitivity test could be used to adjust the coarse grid output, as suggested above.

5. Decouple the model of the channel and the floodplain (to improve model efficiency).

6. Decouple the hydrodynamics from the sediment transport and PCB fate within EFDC (to improve model efficiency).

7. Re-calibrate the model based on suggestions made by reviewers under Question 1, and in consideration of the several instances of bias noted under Question 3. (I don't believe a separate validation is required.)

Response 13-EA-3

4. The suggestion of performing sensitivity analyses for the grid resolution was not implemented because this analysis was estimated to involve a substantial effort that could provide only limited benefit. Simply applying all of the calibration parameters on a more refined grid could lead to "apples to oranges" comparisons. Sub-grid-scale features in the current grid, such as undulations in the bed elevation, have been subsumed into the parameterization of the effective roughness height (z_0). If the resolution of the grid was increased to explicitly represent these features, a reassessment of the appropriate z_0 would be required to produce the same hydrodynamic transport through the new grid. Because data would be required to develop boundary conditions for a laterally segmented grid, the length of a "test reach" would have to span approximately 5 miles, to include locations where monitoring data were collected, as well as water-surface elevations. Obtaining different results from a refined grid of a "test-reach," compared to the results for the same reach from the original grid, would not be meaningful unless there was a basis for concluding that one result was better than the other. A valuable check on the calculations for the full grid of Reaches 5 and 6 is the simulated sedimentation rate in Woods Pond, where estimates of sedimentation rates are available for 12 locations. This check would not be available for simulations performed in only a portion of the domain. Because the Model Working Group concurs that additional lateral segmentation of the river channel is not necessary, EPA believes that the effort required to perform the recommended sensitivity analysis is not justified.

5. The floodplain and channel are effectively decoupled during the majority of the time when in-bank flows are simulated, as a result of logic tests in EFDC, which provide a means of bypassing computations for floodplain cells that are not inundated.

6. The issue of decoupling the hydrodynamics from the sediment transport and PCB fate within EFDC is addressed in the response to General Topic 4.

7. In response to comments from the Reviewers, some parameter changes have been made and the calibration reverified. Revised model results and model-data comparisons, reflecting these changes, are included in the FMD.

Difficult

This step would lead to the greatest model accuracy and robustness, but might take a year or so to accomplish.

8. Restructure the grid to allow more lateral resolution. This would have to include computational efficiency measures so that extended model runs are feasible, as well as a complete model re-calibration.

Response 13-EA-4

8. The issues of lateral grid resolution and code modifications for computational efficiency are discussed in the responses to General Topics 1 and 4.

W. Frank Bohlen

On several occasions discussions with the model group and within the Peer Review Panel made it clear that this required understanding of the variety of transport processes affecting PCB transport in the PSA was less than perfect. Recall the discussions of floodplain dynamics, an area known to represent a significant sink and possible longterm source of PCBs, the continuing debate over sidebank transport, the specification of boundary conditions, and the proper structuring of the sediment bed within the model. Each of these items represent an essential element of the transport model and most will come up again individually in the following review of the model validation effort. Viewed collectively this variety of unknowns makes clear that what EPA and GE are dealing with is a research project rather than simple application of an accepted formulation. With this fact in mind the future role of the model should change to include guidance of monitoring efforts. The need for these additional data should be clear from a scientific standpoint. Their availability would also serve to increase stakeholder confidence in model results.

Monitoring to date has placed primary emphasis on PCB distributions within the sediment column with relatively limited sampling of the water column TSS and PCB concentrations. The latter have placed primary emphasis on flow/TSS relationships during high flow events with sampling at a number of selected transects along the main stem of the river. This monitoring has been supplemented by some few field observations of sidebank erosion and laboratory estimates of bed erodability using SEDFLUME. With the exception of the sidebank observations the majority of the field observations have not been directed at specific processes.

I'd recommend that consideration be given to the extension of these past monitoring efforts to include, for example, the placement of instrument arrays at the Confluence and at the Woods Pond Dam sufficient to provide long term, high frequency (e.g. 3-4 samples/hr), time series observations of water temperature and TSS at the mid-point of the low flow water column. These measurements would be supplemented by monthly sampling of concurrent PCB concentrations. All instrument observations could be telemetered to a central station permitting conditional sampling as unusual flow/transport conditions occur. In addition to the upstream and downstream stations in the PSA consideration should be given to the placement of one or more instrument arrays at selected sites adjoining the flood plain. Again these relatively high frequency data should be supplemented by lower frequency drawn water sampling of concurrent PCB concentrations. This latter sampling might occur on a monthly basis and aperiodically during particular rainfall/runoff events. This combination is intended to significantly increase our understanding of flood plain transport processes and their temporal (including seasonal) variability. These observations would also take allow us to take full advantage of the ongoing remedial efforts by providing quantitative data detailing effect at a number of locations. Such data would seemingly be of value in future remedial planning.

The above observations should be supplemented by a variety of other process studies such as the continuing survey of selected portions of the side banks and sequential bathymetric survey of river channel transects or detailed cross-channel velocity/flux measurements to establish the adequacy of the model grid. Model results would be used to specify siting as well as the need for continuing observations. This close coupling between models and monitoring would be of clear benefit to the long terms goals of this effort.

Beyond these technical issues, the reports provided this reviewer remain exceedingly difficult to read. I understand all of the reasons why but cannot believe that the project would not benefit from a clearer and more concise document. In this validation report there is entirely too much use of “reasonable agreement” and the like with often insufficient demonstration. The Executive Summary is too general and does little to build confidence in this modeling effort and its subsequent application. Questions regarding many of the key points of the model require going back to previous documents that were themselves obscure and it’s often difficult to figure out just what is being presented in the report figures (e.g. was that streamflow instantaneous, daily average or monthly average ? Are the TSS values a vertical average? Over what period of time? Is it legitimate to compare longer term model results to shorter term data?). Too many discussions end prematurely before any attempt is made to explain observed differences or discrepancies (see pg. 4-90/91 Vol.1 discussion of Event 10. Why the underestimation?) Many of these questions might be resolved by a search of our voluminous file but who but the most dedicated would be expected to do it?

I’d recommend, now that the major components of this exercise are in place, that a technical writer be charged with the preparation of a single document describing the model and the resulting runs written for the stakeholder community. This document would include all major features of the model and results with key supporting figures, references and an index. I’d consider this a high priority.

Response 13-FB-1

EPA appreciates the Reviewer’s thoughts regarding additional data collection efforts that could be conducted for the modeling study. Nonetheless, EPA believes that no further additional study is necessary for the model framework to adequately achieve the goal of the modeling study. A massive amount of data has been collected on this site, both as part of the original model framework design and in response to Reviewer comments from the two previous Peer Reviews. These data were sufficient for the implementation of a model calibration approach that ensured that: (1) each individual transport and fate process was correctly parameterized; and (2) the relative importance of each process is appropriately represented. The resulting model Validation provides a further demonstration of model performance and robustness.

EPA agrees that the major components of the modeling study are now in place, and EPA has responded to the Reviewer's comment regarding the preparation of a single document that describes the components of the entire modeling study. The FMD was prepared to summarize the enormous effort resulting from 8 years of interpretation, analysis, and evaluation that occurred in support of the modeling study for the Housatonic River, incorporating input from GE and its consultants as part of the Model Working Group established under the Consent Decree, and in response to the Peer Reviews. The FMD finalizes the outcome of the modeling study and draws upon the information contained in the following documents:

- Draft Model Framework Design (MFD) document and Quality Assurance Project Plan (QAPP) (October 2000)
- Responsiveness Summary to the Peer Review of the MFD and QAPP (June 2002)
- Final MFD (April 2004)
- Model Calibration Report (MCR) (December 2004)
- Responsiveness Summary to the Peer Review of the MCR (January 2006)
- Model Validation Report (March 2006)
- Responsiveness Summary to the Peer Review of the MVR (November 2006)
- Other materials prepared in response to the Peer Reviews that are included in the Peer Review record
- The RCRA Facility Investigation (RFI) Report prepared by GE (September 2003)

In general, comparisons with observed discharge indicates that the watershed model (HSPF) provides accurate simulation of the factors governing stream flow volumes to and through the PSA. These comparisons also suggest that the model is able to reproduce the timing of flow events. This matter of timing is an important factor if the data are to be used to calculate velocity and ultimately boundary shear, as they are in this study and will be in the upcoming Corrective Measures Study. The actual timing of stage/discharge at each section of the study area in large part determines the magnitude of the horizontal pressure gradient which affects speeds, turbulence intensity and boundary shear. These are the principal factors governing sediment/PCB transport both in the water column and across the sediment-water interface.

Given the importance of timing it is disappointing that the report provides so little detailed information on this factor and its effects. There is abundant reference to the model's ability to accurately reproduce event timing but these statements appear to be referring to timing in the most general sense. i.e. a precipitation event induces an increase in streamflow over a time similar to that observed. These references to "event timing" leave the question of the adequacy of the model simulation of timing with respect to flow velocities unanswered. Examination of several of the figures (see Fig. 6.2-5 (attached) e.g.) shows substantial differences between measured and modeled actual flow and stage timing at several stations along the PSA. In this figure, it would appear that the modeled speeds produced by the associated pressure gradients should be less than the measured and that the turbulence induced by adverse pressure gradients would also be reduced. Actual estimation is, of course, complicated by the morphology of the region with tributary inflows and/or backwater flows and storage complicating the simple stage discharge relationship. This factor may be the reason for example that the simulated flows at Holmes Road are higher than the measured despite a higher measured stage relative to that simulated.

Response 13-FB-2

It appears that the Reviewer's assessment that the watershed model, HSPF, accurately predicts the magnitude and timing of flows, correctly recognizes the performance that can be expected from a watershed model. Watershed models are not expected to match measured hourly flows in point-in-time comparisons, even though HSPF does an excellent job of reproducing the frequency distribution of measured hourly flows at both Coltsville and Great Barrington. The performance of watershed models typically improves with increasing averaging periods, which is why the performance measures become more stringent for longer averaging periods. Realistic expectations of HSPF performance on short time scales (e.g., hourly) are important in interpreting EFDC simulations of storm events because of the linkage between HSPF and EFDC. In the simulation of the October 2003 storm event referenced by the Reviewer (MVR Figure 6.2-5), approximately one-third of the flow passing Woods Pond Footbridge is based on HSPF's simulated tributary inflows and direct runoff between the confluence and Woods Pond Footbridge. The discussion of this simulation recognizes the sensitivity to HSPF results, particularly because the double peak in the hydrograph indicates that two closely spaced events were simulated.

Regressions between EFDC simulated and measured flows for New Lenox Road and Woods Pond Footbridge result in coefficients of determination (r^2) of approximately 0.8, with slopes of 0.92 and 1.01. Because these point-in-time comparisons reflect any phase shifts between simulated and measured flows, it is concluded that EFDC does not suffer from timing problems to a degree that is of concern or that is unexpected given the inputs from HSPF.

Does the presence and movement of ice seasonally contribute to sediment erosion in the PSA? Within the channels or along the flood plain?

Response 13-FB-3

In the past 8 years that EPA has worked on the Rest of River, the presence and movement of ice has not been a major factor contributing to sediment erosion in

> the PSA. In addition, in the design of the ½-Mile Removal and 1½-Mile Removal Actions, ice was not considered to be a major design issue. The role of ice influences may be given further consideration during the design of any remedial action for the Rest of River.

The TSS data set appears to be moderately robust although I worry that too much emphasis might have been placed on storm conditions. The data also appear to be primarily point measurements with some representing integrated measurements over the vertical. The absence of time series data prevents detailing of processes such as the time scales of resuspension and deposition during rising or falling stage with particular emphasis on the onset of resuspension during relatively low flow conditions. The influence of such low level but persistent resuspension and transport on PCB fluxes is largely ignored in the present study which places primary emphasis on storm events. Absent the low flow details its difficult to assess the role of these processes in the long term. Such assessments will be a subject of study in the upcoming CMS. As noted above I'd recommend immediate initiation of a monitoring program designed to provide time series observations of TSS concentrations at a number of selected stations throughout the PSA.

Response 13-FB-4

> The emphasis given to storm conditions is warranted because it is under these conditions that the majority of solids and PCB transport occurs. The storm event sampling program did provide time series measurements of water-column solids concentrations; however, the Reviewer may be considering a finer sampling frequency than the 1-hour interval typical of the storm event sampling. The subject of low-level persistent resuspension during low flow was raised by a Reviewer during the Model Calibration Peer Review, and EPA's response was that it was unclear how PCBs and fine-grained solids could accumulate in areas subject to persistent resuspension at low-flow conditions. The Reviewer's recommendation for a monitoring program is addressed as part of Response 13-FB-1 above.

Even with these process questions resolved there remains the issue of run-time. It seems clear that the model as presently configured requires entirely too much time for the completion of a single run to be useful within timely evaluations of a significant number of remedial schemes. We probably knew this several years ago and should have been more sensitive to the need to develop alternative formulations. A number of these, including the separation of the hydrodynamics from the transport estimates and subsequent FCM evaluations were previously mentioned. It is now necessary and possible to go further. Using the experience gained from "whole model" runs it should now be possible to develop a number of synthetic hydrographs detailing streamflow, stage, and TSS concentrations at the upstream boundary and each of the primary tributary streams. This would eliminate the need to run HSPF on a regular basis. EFDC is a ponderous model and can be streamlined now that we have a better idea of the relative importance of the governing variables. John Hamrick should be charged with this task (no more than 6 months) as soon as possible. As part of this streamlining the model grid characteristics should be carefully reviewed, again using what has been learned about the relative importance of each of the domain regions (sidebanks, backwaters, floodplains etc.). My sense of the present grid is that it underspecifies the channel region and overspecifies the floodplain. The latter could almost be treated as a box with fluxes simply proportional to stage which could be specified along its margin by EFDC. If in time more detail of the interior of the plain is needed for remedial purposes consideration might be given to replacing the high resolution grid on the plain while placing a box in some other area (channel sidebanks?). Finally, I'd consider eliminating the FCM in favor of a parametric (flow, TSS concentrations?) approximation of body burden uptake based on the results of the complete model runs.

Response 13-FB-5

EPA disagrees with the Reviewer's comment that the run time prevents the model framework from being "useful within timely evaluations of a significant number of remedial schemes." As discussed in the General Response to General Topic 4, EPA believes that a considerable number of simulations can be executed concurrently, making the model a useful tool for the CMS. EPA disagrees with the Reviewer's suggestions to minimize the roles of the watershed and bioaccumulation models (HSPF and FCM, respectively), since these models have significantly shorter run times than EFDC and little time savings would be gained while considerable rigor would be lost. The Reviewer's suggestion to trade existing grid resolution in the floodplain for additional resolution in the river channel does not appear to reflect the negligible impact that such a change would have on run times. These factors are discussed in FMD Appendix B.1, and in the response to General Topic 3.

In short, what I'm thinking about is the development of an supplementary modeling scheme for use in the CMS. The complete model would serve as a guide assisting in the development of a series of simpler, more efficient but less comprehensive, formulations that would be directed at particular remedial schemes. The complete model framework would remain in place providing guidance regarding the need for and type of data to supplement model formulations for both calibration and verification purposes but would not be run as frequently as the supplementary schema. The alternative might be to turn to a different series of models entirely. This is not recommended without good reasons that I don't have at the moment.

Response 13-FB-6

GE and its consultants will have the option of using techniques suggested by the Reviewer, or other methods to screen scenarios before running the full model. Such options will be discussed by the Model Working Group.

Douglas Endicott

It is unfortunate that the peer review process was conducted in more of a confrontational, as opposed to a collaborative manner. The scientific process is not well served by the former approach.

I think more would have been accomplished in the peer review sessions, had a freer dialog and more open exchange of information and ideas between the various parties been possible.

Because of the constraints of the Consent Decree, this was not allowed. As a result, the relationship between the modeling team and the peer review panel was sometimes adversarial, and our recommendations rebutted inappropriately (from my view). Although this may be as the lawyers intended, it also resulted in some lingering issues that have not been appropriately resolved through the entire peer review process.

Response 13-DE-1

EPA does not believe that the Peer Review process was designed or conducted in a manner to promote or result in a confrontational or adversarial relationship between the parties and the Panel. Nor does EPA believe that any Peer Review recommendation was rebutted in an inappropriate manner. The Peer Review Process outlined in the Consent Decree was developed with deliberate thought to promote an unbiased review of the work products without undue influence from any party to the extent possible.

Outlined below are the steps I believe are necessary to make the EPA models the best possible tools to accomplish the goals of the modeling study for the Housatonic River:

- Revise the MVR to incorporate supplemental material (*EPA Response to Questions* document , Document Overview presentations, etc.) and remove provisional results (mass balance diagrams, downstream model w/o boulders, …) that have been superceded since the MVR was released.
- Make a number of near-term corrections to EFDC, which can be addressed/resolved within a ~1 yr time frame)
    - Correct parameterization errors identified in EFDC model:

      Reduce thickness of surficial sediment bed layers;

      Parameterize vertical mixing rates as functions of benthos density & vertical position in sediment bed;

      Increase diffusive PCB fluxes by calibration and parameterize spatially as functions of benthos density;

      Reanalyze SEDFLUME data using resuspension parameters constrained by literature (e.g., shear stress exponent n=2.6±0.3) and recalibrate deposition rates. Lick, Ziegler and coworkers have published much guidance on the specifics of parameterizing sediment resuspension[1], some of which the modeling team has chosen to ignore;

---

[1] See, for example: A Quantitative Framework for Evaluating Contaminated Sediment Sites. SETAC 20th Annual Meeting, Philadelphia, PA, November 14-18, 1999.

- Revise TSS and PCB boundary conditions to correct bias evident in low flow range.

- Test sensitivity of EFDC hydrodynamics, sediment transport and PCB transport simulations to alternative grid resolutions in a "test reach" of the PSA. In all prior SEDZL applications I am aware of, at least a 2-dimensional model was used for hydrodynamics and sediment transport. In river systems, this was usually a vertically-integrated model. Furthermore, on rivers that bend as much as the Housatonic, a curvilinear grid has commonly been applied. Since the EPA modeling team has elected not to follow these conventions, they should at least demonstrate via numerical testing that their primarily 1-dimensional model of the Housatonic River channel produces comparable results for sediment and PCB transport.

- Investigate methods to economize on hydrodynamic and sediment transport simulations (e.g., using steady state design storms and flow-duration statistics, such as the Saginaw River SEDZL application[2])

- Consider using the calibrated and validated EFDC model to build a simpler box model of the river, if the two models can be shown to produce comparable results. The simpler model could be particularly useful for forecasting uncertainty. Don Mackay suggested this "second simplified model" approach for modeling hydrophobic contaminants in the Niagara River[3].

- Develop and implement a process of ongoing model validation, using data collected as remediation progresses. It goes with out saying that modelers want to continue modeling… However, the reality is that no model is truly ever "finished". So long as new data are collected, model refinement must be expected and accommodated by managers and decision makers.

Response 13-DE-2

- In response to comments from several Reviewers, the Final Model Documentation Report (FMD) has been prepared to synthesize information presented in the MFD and the Model Calibration and Validation reports, and contained in the various Peer Review documents.

- In response to comments from the Reviewers, some parameter changes have been made. Revised model results and model-data comparison, reflecting these changes, are included in the FMD.

---

[2] Cardenas, M. and W. Lick. 1996. Modeling the Transport of Sediments and Hydrophobic Contaminants in the Lower Saginaw River. *J. Great Lakes Res.,* 22(3):669-682**.**

[3] McLachlan, M. and Mackay, D., "A Model of Contaminant Fate in the Niagara River", report prepared for Environment Canada (1987).

- In response to comments from the Peer Reviewers concerning the bed layering in EFDC, sediment profile imaging was conducted in early September 2006 to obtain site-specific information on the structure of the sediment of the Housatonic River. The site-specific information obtained from that effort supported the bed layering used in the Phase 2 Calibration and Validation modeling, so no change was made to the sediment bed layer thickness (see FMD Appendix B.5 for additional information).

- Vertical mixing rates in the sediment bed are parameterized as a function of benthic invertebrate density. During the Phase 2 Calibration, an extensive literature review was conducted when parameterizing the bed layering and mixing. The vertical structure of particle mixing in the bed was revisited following the Model Validation Peer Review, including attention to the issue of the vertical position of benthic organisms in the sediment bed. Additional literature review was conducted as part of this effort. An adjustment was made to the algorithm used in EFDC to specify mixing between bed layers, as described in Appendix B.5 of the FMD.

- EPA disagrees with the Reviewer's opinion that there is a need to increase diffusive PCB fluxes. The recommendation to parameterize the sediment-water mass transfer coefficient as a function of benthos density presumes this to be the dominant factor controlling the mass transfer coefficient, when this is not known to be the case. A more detailed discussion of the parameterization of the sediment-water mass transfer coefficient is provided in Appendix B.10 of the FMD and is discussed further in General Topic 1.

- In response to Reviewers' comments, the exponent of the normalized-excess shear stress, used to compute the cohesive solids erosion rate, was re-evaluated, as described in Appendix B.8 of the FMD. Revised model results and model-data comparisons, reflecting these changes, are included in the FMD.

- In response to Reviewers' comments, additional effort was directed to the estimation of upstream boundary conditions for times when data are not available. Alternate analysis techniques were attempted, including the Autobeale method recommended by this Reviewer. A detailed discussion of the boundary condition estimation techniques is provided in Appendix B.2 of the FMD.

- A response to the recommendation of using a test reach to evaluate the sensitivity of the simulation results to an alternate grid is provided in Response 13-EA-3 above.

- Responses to General Topic 4 address a variety of suggestions for reducing the computational burden associated with the hydrodynamic and sediment transport calculations. It is noted that the Reviewer's suggestion of using steady-state design storms would eliminate the ability of the model to simulate the flow acceleration at the onset of a storm event. The importance of this factor was stressed by Dr. Bohlen in his comment (preceding Response 13-FB-2 above) regarding the importance of reproducing timing of flow events, "The actual timing of stage/discharge at each section of the study area in large part determines the magnitude of the horizontal pressure gradient which affects speeds, turbulence intensity and boundary shear."

- The application of the model in the CMS will be discussed in the Model Working Group and subsequently proposed by GE and reviewed by EPA. Techniques for improving efficiency, such as that described by the Reviewer, may be considered.
- The process and timeline for evaluation and selection of a remedy for the Rest of River is specified in the Reissued RCRA Permit and the Consent Decree. Within this process, the Model Working Group will have continued discussions regarding the use of the model. Any consideration to conduct further modeling will be given within that framework.

Marcelo H. Garcia

I do not think that the model is ready to accomplish the goals of the modeling study. If the model is to be used to simulate future conditions, first is has to adequately simulate the existing conditions throughout the river and its floodplain.

Response 13-MG-1

EPA strongly disagrees that the model is not ready to accomplish the goal of the modeling study, and that it does not adequately simulate the existing conditions in the Housatonic River and floodplain. The model framework was used to successfully predict average PCB concentrations in water, sediment, floodplain soil, and aquatic biota. In consideration of model performance, strengths and limitations, the framework was judged to reasonably account for the important processes that control PCB fate, transport, and bioaccumulation in the river system at the relevant scales. The evaluations of model performance also demonstrate that the model framework appropriately characterizes the balance of fate processes affecting PCB concentrations and exposures over time. See Responsiveness Summary General Topics 1 and 2.

Frank Gobas

The modelers should improve the calibration of the model by lengthening the model calibration period. This most likely involves a continuation of monitoring the response of PCB concentrations in the River to the remediation efforts that have recently taken place. Over time, a more definite change in concentration may take place, which can be used to better calibrate the model.

Response 13-FG-1

EPA does not believe that it is necessary or desirable to continue modeling efforts. Remediation above the upstream boundary condition will not be completed for a number of years, thus the collection of monitoring data to characterize the full effects of the remediation would take years. EPA believes that this would result in an unacceptable delay in the implementation of the Rest of River remedy decision.

An alternative to the difficult characterization of the current mass of PCBs in the River, is using the model to investigate under which set of model parameters historical sources are the main contributor to the PCB concentrations in the Housatonic River and judge whether these model parameter sets are reasonable. This would avoid having to characterize the actual current mass of PCBs in the River.

Re. Model Objective #3:

Due to the heavy reliance on calibration during model development, the relative PCB concentration data in water, sediments and biota that are calculated by the model are overall consistent with the model predictions of the mean concentrations. The model can therefore be used with confidence to address the relative contributions of current PCB sources to bioaccumulation in target species (i.e. model objective #3). Dr. Connolly argued that the potential underestimation of the overall PCB depuration rate in the River affects estimates of the relative contributions of current PCB sources to bioaccumulation in target species. While this is correct, I do not think that this will have a significant effect on the derivation of current relative sources of PCBs to fish because PCB concentrations did not vary significantly over the time period that calibration was performed.

The application of the model of the model to quantify the historic contributions of various PCB sources to bioaccumulation in target species is dependent on the time dependent capabilities of the environmental fate model, which are more uncertain (see discussions above). The FCM model can be expected to properly estimate the relevant contributions of water and sediment concentrations as sources of PCB bioaccumulation in benthos and fish species from PCB concentrations in water and sediments delivered by the EFDC model.

Response 13-FG-2

EPA agrees with the Reviewer's comment that "The FCM model can be expected to properly estimate the relevant contributions of water and sediment concentrations as sources of PCB bioaccumulation in benthos and fish species from PCB concentrations in water and sediments delivered by the EFDC model." The Reviewer's concern about the ability of EFDC to assess the relative importance of current and historical PCB sources appears to stem from the Reviewer's opinion that the temporal response computed in EFDC is uncertain. Despite wide confidence bands on mean sediment PCB concentrations in

> different years, several Reviewers, including Dr. Gobas, expressed the opinion that the data showed a significant decline over time, and that this decline was underestimated by the EFDC simulation. The results of a statistical analysis conducted in response to Reviewer's concerns as well as previous analyses presented to the Reviewers, indicates that the surface sediment PCB concentrations show no statistically significant change over time, and that the temporal response simulated by EFDC was well within the uncertainty envelope of the data. This information, which is presented in Appendix A.1 of the FMD, should address the Reviewer's concerns regarding the temporal response simulated by EFDC.

A second suggestion is to lengthen the model's calibration period. The calibration period for the current model was too short for an evaluation of model capability. This would involve the continuation of monitoring programs with the objective to develop PCB concentration data over a longer period of time.

Response 13-FG-3

> EPA lengthened the calibration period to 10 ½ years in response to Reviewers' comments following the Peer Review of the Model Calibration Report. At that time the Reviewers suggested that an appropriate calibration period would be between 5 and 10 years. See also Response 13-FG-1.

Continue existing PCB concentration monitoring programs to measure the changes in PCB concentrations over time as a result of the recently completed remediation. Use the data together with calculations of PCB source reductions due to remediation to extend the calibration period and improve the calibration and/or validation of the long term temporal response of the model.

Apply a staged and adaptive approach in the planning of River remediation. Plan to gauge the river's response to remedial efforts at certain locations in the River throughout the River remediation. A PCB concentration monitoring program can detect the effect of remedial actions on PCB concentrations over time and space. These data can then be used in the model to further optimize the model, such that the effects of newly planned remedial efforts can be better estimated.

Response 13-FG-4

> It is expected that GE will continue to conduct sampling for PCBs and TSS at the established monthly monitoring locations that span the areas in which remediation has been or is being conducted. See Response 13-FG-1.
>
> The approach to evaluating and selecting a remedy for the Rest of River is outlined in the Consent Decree and Reissued RCRA Permit. This is the process that EPA and GE will follow in planning the remediation activities. It is expected that any selected remedial strategy other than no action will include a monitoring component.

Wilbert Lick

The present model is inadequate to achieve the goal of the modeling study to simulate future conditions (1) in the absence of remediation and (2) for use in evaluating the effectiveness of remedial alternatives.

Response 13-WL-1

EPA rejects the Reviewer's assertion that the model framework is inadequate. As documented by the analyses and results in the FMD and in previous modeling documents and in individual responses to the other comments in this Responsiveness Summary, EPA believes that the model framework is adequate to achieve the goal of the modeling study. EPA believes the Reviewer's assertion is unfounded because it is the result of flawed logic that disregards site-specific measurements, the model development methods used, and the stated goal of the modeling study. The discussion below presents the technical basis for EPA's conclusion that the model framework is adequate and that the Reviewer's position is incorrect.

First, it should be recognized that the model framework for Housatonic River was developed using site-specific data to the greatest extent practical. Some of the most important site-specific measurements included net sedimentation rates, cohesive sediment erosion properties, sediment profile images, surface sediment PCB concentration trends over time, the characteristics of benthic organisms in surface sediment, PCB partitioning, and the sediment-water mass transfer flux of PCBs. These measurements define acceptable ranges of parameter values for many model processes. In many cases, the processes defined by site-specific data are the same ones the Reviewer characterized as inaccurate or incorrect and in conflict with data. The following two specific examples highlight the Reviewer's flawed logic. One basis the Reviewer uses to assert the model calibration is "inaccurate" is that there is an infinite number of combinations of erosion and net sedimentation rates that can reproduce observed suspended solids concentrations as long as the deposition rate is changed accordingly (see Appendix A, Lick comments, page 3). The Reviewer further stated that "for a predictive model, the values of erosion rate and deposition rate can not both be determined from calibration of the model by use of the suspended solids concentration alone." The flaws in the Reviewer's logic are that the model calibration in fact was not based on suspended solids concentrations alone and that two additional parameterization constraints exist: (1) water column PCB concentrations; and (2) the surface sediment PCB concentration rate of change. These constraints limit the number of plausible model parameterizations to a narrow range because a large fraction of the PCBs present in the water column or sediment bed is associated with particles. When these particles are eroded from or deposited to the bed, water column PCB concentrations change. Erosion and deposition can also cause sediment bed PCB concentrations to change over time. Consequently, the deposition rate cannot be artificially set to an unrealistic value because the model parameterizations must simultaneously satisfy four conditions: (1) water column suspended solids concentrations; (2) net sedimentation rates; (3) water column PCB concentrations; and (4) surface sediment concentration changes over time. As documented in Sections 3 and 4 of the FMD, the model framework successfully simulates water column solids and PCB concentrations while simultaneously reproducing net sedimentation rates and the surface sediment PCB concentration rates of change.

Another basis for the Reviewer's assertion that the model calibration is "inaccurate" (and "incorrect") is that a PCB sediment-water mass transfer flux is used to represent a number of individual processes such as groundwater advection, diffusion, and bioturbation. The Reviewer comments that the model is wrong because each of these processes must be modeled separately with each having a different functional form (see Appendix A, Lick comments, page 6). From this initial assertion, the Reviewer goes on further to say that the PCB sediment-water mass transfer flux is dependent on bioturbation and the existence of well-mixed sediment layers. The flaws in the Reviewer's logic are that none of these assertions is individually or collectively true. Specifically, it should be noted that the PCB sediment-water mass transfer process is used to represent the composite effect of individual processes that cannot be readily measured or otherwise quantified in a practical manner on an individual basis. This means that, regardless of differences in functional form, the effect of each process is so small that they cannot be simulated on an individual basis without introducing new model parameters that cannot be defined from data. Consequently, model accuracy cannot be improved by decomposing the composite process into unmeasureable components. Further, EPA did not assume (and there is no reason to assume) that PCB mass transfer from the sediment bed is solely driven by bioturbation or that a well-mixed layer must exist for PCB mass transfer to occur. Further, the composite effect of PCB mass transfer from the sediment bed is well-defined by site-specific data as documented in FMD Appendix B.10. Despite his critique, the Reviewer nonetheless acknowledges that EPA used the generally accepted modeling approach (i.e., representation of the sediment-water mass transfer flux as a gradient-driven process). Given the site-specific basis for the sediment-water mass transfer process parameterization, the Reviewer's assertions are unfounded.

Second, it should be noted that a fundamental attribute of EPA's model calibration strategy was that each transport and fate process was independently parameterized. The site-specific data were examined to isolate conditions where sediment or PCB transport was controlled by a single, predominant process. This calibration strategy was particularly important to the overall model development effort because it ensured that: (1) each transport and fate process was correctly parameterized; and (2) the relative importance of each process (termed the "balance of fate processes") was appropriately represented. This approach is documented in FMD Section 3.3. Many comments leading to the Reviewer's assertion that the model cannot meet the goal of the modeling study are based on the Reviewer's failure to recognize that the model correctly represents the balance of fate processes. A specific example of the Reviewer's flawed logic in this regard is evident in his assessment of the impact that sediment mixing has on PCB fate. The Reviewer falsely assumed that: (1) mixing is exclusively driven by benthic organisms; and (2) that the mass transfer flux of PCBs should necessarily change as the benthic biomass changes.

Third, the Reviewer neglected key elements of the stated goal of the modeling study, particularly the intended spatial and temporal scales of model use. In his analysis of model performance and recommendations, the Reviewer focused on very fine spatial and temporal scale processes and did not recognize that the detail he recommends is not necessary to meet the study goal. The Reviewer repeatedly focused on very fine spatial and temporal scales and other factors that cannot be accurately measured in the Housatonic River system. In contrast, EPA focused on those processes and effects that could be measured. At the intended scale of use, processes such as PCB release by diffusion and even groundwater advection are not individually important. As previously noted, EPA represented these processes in a composite manner because only the composite effect of the individual processes is measurable at the intended scale of model use. As documented in FMD Appendix B.10, EPA has correctly represented the composite effect of processes contributing to the sediment-water mass transfer flux of PCB at the intended spatial and temporal scales of model use.

Further, the Reviewer asserted that use of equilibrium partitioning (EqP) in the model is necessarily wrong and has a "major effect" on PCB transport (see Appendix A, Lick comments, pages 9 and 12). However, the Reviewer did not examine if his initial premise was true that the use of EqP has a major effect on PCB transport in the Housatonic River before concluding whether the use of EqP is wrong. For the specific case of the Housatonic River, EPA's calculations demonstrate that use of either kinetic or EqP partitioning approaches are expected to differ by a mere 5 to 10% as documented in FMD Appendix B.9. A potential 5 to 10% difference in model results is less than the measurement error in PCB sorption rates or partition coefficients. On this basis, EPA concluded that EqP is appropriate for this study while kinetic approaches are unnecessarily complex and not demonstrably more accurate.

As applied to the Housatonic, sensitivity and uncertainty analyses are asking the wrong questions. They do not question whether the basic processes are formulated correctly. As an example, equilibrium partitioning is assumed. Sensitivity and uncertainty analyses never question this assumption, never demonstrate that it is an inaccurate assumption, nor do they propose a suitable reaction rate.

Response 13-WL-2

The central point of the Reviewer's comment is concern about the description of processes or formulations included in EFDC. Responses to the Reviewer's concerns about processes are addressed in the responses to General Topic 1. The applicability of the specific formulation mentioned by the Reviewer, equilibrium partitioning, is discussed in detail in Appendix B.9 of the FMD. EPA strongly disagrees with the Reviewer's opinion that equilibrium partitioning "is an inaccurate assumption." As discussed in Appendix B.9 of the FMD, equilibrium partitioning is an approximation that is clearly valid for the Housatonic River. Under relevant conditions for PCB remobilization, differences between the results of the equilibrium partitioning assumption and the most extreme case of slow desorption kinetics are less than 10% in calculated PCB concentrations on suspended particles. This level of uncertainty is clearly tolerable considering other, more significant sources of uncertainty (including measurement error) and the computational burden associated with representing sorption and desorption kinetics.

These suggested modifications to the model are relatively simple (except possibly for the re-gridding) and should be able to be accomplished in a year.

Response 13-WL-3

See the body of the Responsiveness Summary for responses to the suggested modifications. EPA has, where appropriate, evaluated the suggested modifications.

E. John List

The Peer Review Committee (PRC) has spent much time reviewing the results of the modeling exercises and analyzing the modeling performance. It has devoted a significant level of effort to suggestions as to how to improve the model performance and enhance its predictive capabilities in its application to remedial scenarios. However, to my mind the PRC has also perhaps focused excessively on model omissions and shortcomings and, in the process, has tended to overlook the very real benefits that have come from the modeling exercise, benefits that are substantial, irrespective of the perceived model shortcomings.

First, I believe that the most valuable aspect of the modeling exercise has been the discipline that has been imposed to the assessment of a detailed Housatonic River mass flux balance for the PCBs, total suspended solids (TSS) and water. The key factor in the assessment of any remedial strategy is going to be how it affects these mass fluxes. The modeling exercise enabled a detailed assessment of the relative order of magnitude of the mass fluxes in the various river system components and how these would change under different source flux hypotheses. Although these absolute fluxes may be somewhat in error due to the model shortcomings (to be discussed below), the relative order of magnitude under different hypotheses enables the key leverage points of control for the PCB fluxes to be identified. The available flux data appear to indicate that, although the modeling may have some problems, at least the predicted orders of magnitude appear to be in the range of available data. This suggests that the flux distributions, as described in Figures 1-5 and 1-6, and Figures 1-9 through 1-15, of the document "EPS Responses to Questions from Model Validation Document Overview Meeting", are most probably "in the ball park" and, when properly validated, can be used to develop potential remedial strategies.

Response 13-JL-1

EPA agrees with the Reviewer's observation that some very real benefits have come from the modeling exercise and the input from the Panel. EPA believes that the result of the modeling study is a model framework that successfully predicts average PCB concentrations in water, sediment, floodplain soil, and aquatic biota. In consideration of model performance, strengths and limitations, the framework is judged to reasonably account for the important processes that control PCB fate, transport, and bioaccumulation in the river system at the relevant scales. The evaluations of model performance also demonstrate that the model framework appropriately characterizes the balance of fate processes affecting PCB concentrations and exposures over time. Therefore, the model framework provides a tool that can be used as one line of evidence with other information, including data analyses and the risk assessments, to achieve the goal of the modeling study.

In my professional opinion these two items are really key issues in the fate and transport analysis for the PCB. If neither phenomenon can be explained then how can any model hope to represent what is now occurring or, even more important from the point of view of this project, what will occur in the future. The second item I suspect may be a result of not including wind stresses and wind-induced mixing in Woods Pond. Wind waves continually stir up the sediment and when the wind dies the fine sediment (which apparently is high in PCB) settles last and stays there until the next set of wind waves occurs. The wind waves result in a preferential winnowing of fine material into the surficial sediments of the pond.

Response 13-JL-2

The initial portion of the Reviewer's comment relates to concerns that EFDC does not simulate variability in sediment PCB concentrations consistent with the variability observed across individual samples. The Reviewer continues to express this concern, even though EPA and the modeling team have described that this should be expected because sediment PCB concentrations from 0.25 mile, or more, sections of the river were averaged to develop initial conditions for sediment PCBs. More detailed discussion is provided in the responses to General Topics 2 and 7.

The latter portion of the Reviewer's comment discusses potential effects of winds on sediment resuspension and transport processes, which were evaluated by the modeling team. The topography surrounding the Housatonic River limits the wind fetch, which is an important parameter affecting wind-driven circulation and bottom shear stresses from wind-waves. The wooded hills that rise in elevation on both sides of Woods Pond limit the wind fetch in the section of Reaches 5 and 6 with the widest exposed surface area. The effect of wind on sediment resuspension was investigated by specifying a time series of wind speed and direction in EFDC. Because of the sheltering effects of the surrounding topography, wind effects were negligible. The Reviewer's comment regarding the potential role of winds on sediment resuspension and transport was not supported by the evaluation performed with EFDC.

With respect to the PCB modeling the results are so tenuous (see Figure 6.2-49) that it is not possible to draw any really solid conclusion with respect to bias. However, an important point is demonstrated by Figure 6.2-50. This figure indicates that the sediment concentrations of PCB in the lower reaches of the project study area have declined somewhat since 1990 (i.e., since remediation activity started). Furthermore, there appears to be a similar decline in the water column concentrations of PCB, as is somewhat evident in the lower panel of Figures 6.2-44. (It is noted that there is no statistical measure of the significance of the decline in either the sediment or the water column, but the data are very suggestive). However, the modeling does not give any indication that any similar rate of decline has occurred (see Figure 6.2-50). In fact the comparison between the data and the model in this figure is strongly suggestive that the model does have a bias. The bias of the model is further reinforced by the results of the hypothetical remedial scenarios plotted in the un-numbered figure on page 56 of the EPA Response to Questions from Model Validation (Document DCN: GE-061406-ADFI), where, despite a reduction in the flux of PCB from the sediments in the upper river reaches, there is no significant change in projected sediment concentration of PCBs in the lower reaches. Thus the model does not match the reduction in sediment concentration that has occurred subsequent to actual post-upstream remediation, and projects no reduction as a result of an hypothetical reduction in the upstream PCB flux, as in the hypothetical examples. This problem needs to be understood and/or fixed before the model is used in real applications.

Response 13-JL-3

The Reviewer states that the data shown in Figures 6.2-49 and 6.2-50 are suggestive of statistically significant declines in sediment PCB concentrations. The data were presented as means ± 2 standard errors, which approximate the 95% confidence interval of the mean. The wide range in the confidence interval of the mean, and the small number of samples in earlier years was cited by the modeling team as factors that complicated characterization of trends in the data and comparisons with trends computed by the model. In response to comments from Reviewers, a more detailed statistical analysis has been performed in addition to the analyses previously conducted and presented to the Reviewers. These analyses indicate that there is no statistically significant decline in sediment PCB concentrations. The result of this analysis does not support the basis for the Reviewer’s conclusion that the model-data comparisons indicate a bias in the model.

The results of the example simulations show a gradual decline in sediment PCB concentrations in the sediments of downstream reaches in the base case and simulations 1 and 2, which include elimination of upstream sources. The differences among the rates of decline in sediment PCB concentrations in the base case and example simulations 1 and 2 are small, which is consistent with the observation that conditions in the river are changing slowly.

The Reviewer’s statement that remediation began in 1990 is incorrect. With the exception of limited activity associated with Building 68, the remediation in the ½-Mile Reach began in October 1999, and remediation in the 1½-Mile Reach began in 2002. The Reviewer’s evaluation of the results of the example scenarios in comparison to data is based on an incorrect assumption that the data reflect 9 years of remediation activities.

The primary difficulty here is that the processes that result in the observed concentration distributions of PCBs in the sediment are not known and therefore cannot be included in the model. Second, the geometry of the system is probably too complex to apply properly the sediment transport theory that is known. However, these difficulties should not prevent the development of adequate empirically-based relationships between observed PCB concentrations and fluxes and the water and sediment fluxes. The sediment transport modeling attempts to use a fine scale theory that is applicable to a uniform flow and applies it over an entire river cross-section as if the river had this uniform flow. Many careful observations in the field, of PCB concentrations and fluxes together with stream flow and sediment fluxes, would have provided empirical relationships that would likely have proven much more accurate, useful and predictable. There is nothing wrong with such empirical relationships when the basic theory is unknown or too complex, and in fact EPA and the USGS already use such relationships (see the two websites referenced above).

Response 13-JL-4

Responses to the Reviewer’s initial comment regarding processes are provided in the responses to General Topic 1. The remaining comment regarding an empirical approach that could have been developed, had different data been collected, does not represent an actionable recommendation for this modeling effort.

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.

Burkhard, L.P. 1998. Comparison of two models for predicting bioaccumulation of hydrophobic organic chemicals in a Great Lakes food web. *Environmental Toxicology and Chemistry* 17(3):383-392.

Burkhard, L.P. 2000. Estimating dissolved organic carbon partition coefficients for nonionic organic chemicals. *Environmental Science and Technology* 34(22):4663-4668.

Chapra, S.C. 1997. *Surface Water Quality Modeling.* McGraw Hill Publishing, New York, NY.

Di Toro, D.M. 1985. A particle interaction model of reversible organic chemical sorption. *Chemosphere* 14(10):1503-1538.

Eadie, B., N. Morehead, and P. Landrum. 1990. Three-phase partitioning of hydrophobic organic compounds in Great Lakes waters. *Chemosphere* 20(1-2):161-178.

Eadie, B., N. Morehead, V. Klump, and P. Landrum. 1992. Distribution of hydrophobic organic compounds between dissolved and particulate organic matter in Green Bay waters. *Journal of Great Lakes Research* 18(1):91-97.

EPA (U.S. Environmental Protection Agency). 1994. *Validation Strategy for the Integrated Exposure Uptake Biokinetic Model for Lead in Children*. Office of Solid Waste and Emergency Response, U.S. Environmental Protection Agency, Washington, DC. EPA 540/R-94-039 PB94-963504 OSWER #9285.7-21. December 1994.

EPA (U.S. Environmental Protection Agency). 2000. *Phase 2 Report, Further Site Characterization and Analysis, Volume 2D, Revised Baseline Modeling Report, Hudson River PCBs Reassessment RI/FS*. Prepared for U.S. EPA and U.S. Army Corps of Engineers Kansas City District by Limno-Tech, Inc., Menzie-Cura and Associates, Inc., and TetraTech, Inc. January 2000.

EPA (U.S. Environmental Protection Agency). 2006. *EPA Response to Questions from the Model Validation Document Overview Meeting.* DCN GE-061406-ADFI. June 2006.

Gailani, J., C.K. Ziegler, and W. Lick. 1991. The transport of sediments in the Fox River. *J. Great Lakes Res.* 17:479-494.

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

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

Jepsen, R., S. Borglin, W. Lick, and D.L. Swackhammer. 1995. Parameters affecting the adsorption of hexachlorobenzene to natural sediments. *Environmental Toxicology and Chemistry* 14(9):1487-1497.

Karickhoff, S.W., D.S. Brown, and T.A. Scott. 1979. Sorption of hydrophobic pollutants on natural sediments. *Water Research* 13:241-248.

Kashian, D.R., Prusha, B., and Clements, W.H. 2004. Influence of total organic carbon and UV-B radiation on zinc toxicity and bioaccumulation in aquatic communities. Environmental Science and Technology, 38(23):6371-6376.

Kleijnen, J.P.C. 1999. Validation of models: statistical techniques and data availability. In: P.A. Farrington, H.B. Nembhard, D.T. Sturrock, and G.W. Evans (Eds), *Proceedings of the 1999 Winter Simulation Conference*. pp. 647-654.

Knighton, D. 1998. *Fluvial Forms and Processes: A New Perspective.* Arnold Publishers, London, U.K. 383 p.

Leopold, L., M. Wohlman, and J. Miller. 1964. *Fluvial Processes in Geomorphology*. Dover Publications, Inc., Mineola, NY. 522 p.

Lick, W., J. Lick, and C.K. Ziegler. 1994. The resuspension and transport of fine-grained sediments in Lake Erie. *Journal of Great Lakes Research* 20(4):599-612.

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.

McNeil, J., C. Taylor, and W. Lick. 1996. Measurements of erosion of undisturbed bottom sediments with depth. *J. Hyd. Engrg.* 122(6):316-324.

QEA (Quantitative Environmental Analysis, LLC). 1999. *PCBs in the Upper Hudson River.* Prepared for General Electric, Albany, NY. May 1999. Amended July 1999.

QEA (Quantitative Environmental Analysis, LLC). 2001. *A Model of PCB Bioaccumulation in the Lower Fox River and Green Bay: GBFood.* Prepared for ThermoRetec, St. Paul, MN. June 2001.

QEA (Quantitative Environmental Analysis, LLC). 2006. Housatonic River, General Electric Company's Comments on EPA's Model Validation Report, MVR Peer Review Meeting, June 28, 2006.

Richards, R.P. 1998. *Estimation of Pollutant Loads in Rivers and Streams: A Guidance Document for NPS Programs*. Water Quality Laboratory, Heidelberg College, Tiffin, OH. Project report prepared under Grant X998397-01-0, U.S. Environmental Protection Agency, Region 8, Denver. 108 p.

Schwartz, S. 2000. Quality Assurance of Exposure Models for Environmental Risk Assessment of Substances. Institute of Environmental Systems Research, Department of Mathematics and Computer Science, University of Osnabrück, Germany. Doctoral thesis, June 2000.

Schwarzenbach, R.P., P.M. Geschwend, and D.M. Imboden. 1993. *Environmental Organic Chemistry*. Wiley-Interscience, New York.

Shelly, A., D. Ford, and B. Beck. 2000. Report of a Workshop on Quality Assurance of Environmental Models, Seattle, September 7 - 10 1999. National Research Center for Statistics and the Environment (NRCSE) and United States Environmental Protection Agency (USEPA). NRCSE Technical Report Series No. 042. April 19, 2000.

Steuer, J., S. Jaeger, and D. Patterson. 1995. *A Deterministic PCB Transport Model for the Lower Fox River between Lake Winnebago and DePere, Wisconsin*. Wisconsin Department of Natural Resources, Publ. WR 389-95.

Thomann, R., J. Mueller, R. Winfield, and C. Huang. 1991. Model of fate and accumulation of PCB homologs in Hudson Estuary. *Journal of Environmental Engineering* 117(2):161-178.

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

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.

Vuksanovic, V., R. De Smelt, and S. Van Meerbeeck. 1996. Transport of polychlorinated biphenyls (PCB) in the Scheldt Estuary simulated with the water quality model WASP. *Journal of Hydrology*, vol. 174, no. 1/2, p. 1.

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

WESTON (Weston Solutions, Inc.). 2002. *Responsiveness Summary to 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. DCN GE-053102-ABAQ.

WESTON (Weston Solutions, Inc.). 2004a. *Modeling Framework Design: Modeling Study of PCB Contamination in the Housatonic River*. Prepared for the 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.

WESTON (Weston Solutions, Inc.). 2006a. *Responsiveness Summary to the Peer Review of 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-080105-ACUD.

WESTON (Weston Solutions, Inc.). 2006b. *Model Validation: 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-030706-ADBR.

WDNR (Wisconsin Department of Natural Resources). 2001. *Development and Application of a PCB Transport Model for the Lower Fox River*. Wisconsin Department of Natural Resources, Madison, WI. June 15, 2001.

Woodlot (Woodlot Alternatives, Inc.). 2002a. *Erosion and Accretion Mapping Meander Study, November 2001-June 2002, for Housatonic River Primary Study Area*. Prepared for U.S. Environmental Protection Agency. DCN: GE-070302-ABCD. July 2002.

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

APPENDIX A

PEER REVIEWER COMMENTS

APPENDIX A.1

COMMENTS OF ERIC ADAMS

Final Comments on Model Validation Report
E. Adams, July 2006

Question 1 Considering the changes implemented in the Phase 2 Calibration, does the model reasonably account for the relevant processes affecting PCB fate, transport and bioaccumulation in the Housatonic River to a degree consistent with achieving the goal of the modeling study?

▶*Sorption kinetics*

The model assumes that three chemical phases of PCBs are in equilibrium. As EPA points out, this is a common modeling assumption, but I believe it is only reasonable within a stationary sediment bed, and not within the water column during sediment resuspension.

Following Wu and Gschwend (1986), and focusing only on sorption and desorption among two phases (sorbed PCB with concentration C_s and dissolved PCB with concentration C_d)

$$\frac{dC_s}{dt} = k_1 K_p C_d - k_1 C_s \qquad (1)$$
$$\frac{dC_d}{dt} = k_2 C_s / K_p - k_2 C_d$$

where k_1 and k_2 are rate constants constrained by

$$k_1 = k_2 / \rho K_p \qquad (2)$$

where ρ is the solid-water phase ratio $\rho = (1-\phi)\rho_s / \phi$, ρ_s is the solid density, ϕ is porosity and K_p is the partition coefficient. For a stationary sediment bed, ρ is about 1, so for hydrophobic contaminants (large K_p), most of the contaminant is sorbed; hence the processes of sorption and desorption cause C_d to vary in the range $0 \le C_d \le C_s / K_p$, whereas C_s remains nearly constant. The characteristic time for sorption/desorption for highly particle-reactive species in a stationary sediment bed ($\rho K_p >> 1$) is thus k_2^{-1}.

Wu and Gschwend (1986) describe the sorption/desorption process as one of molecular diffusion through particles (or particle aggregates). For large K_p and an assumed intra-aggregate porosity, their effective (retarded) diffusion coefficient is given by

$$D_{eff} \cong \frac{0.2 D_m}{K_p \rho_s} \qquad (3)$$

where D_m is the molecular diffusivity of PCBs. Wu and Gschwend (1988) fit a first-order sorption kinetics model to the radial diffusion model over the (initial) time during which half of the sorption/desorption takes place. Comparing the two models, and for $\rho K_p >> 1$

$$k_2^{-1} \sim \frac{(R/\rho K_p)^2}{D_{eff}} \tag{4}$$

where R is a characteristic aggregate radius. It can be seen that the time scale (k_2^{-1}) is proportional to K_p^{-1}. In other words, for highly particle reactive species, little of the initially sorbed mass is exchanged during desorption and that which is, is lost from the outside of the particle (over an effective length of order $R/\rho K_p$). Even for a large R (0.01 cm) and $K_p = 10^5$ cm^3/g, the time scale (k_2^{-1}) is less than one second, suggesting that desorption is very fast, and that the assumption of ***equilibrium partitioning is probably acceptable within the sediment bed***. See also following discussion under diffusive exchange.

During resuspension, where the sediment concentration in the water column (TSS) is small, equilibrium partitioning would require that most contamination become desorbed. Hence diffusion must take place through the entire particle, giving a time scale (k_1^{-1})

$$k_1^{-1} \sim \frac{R^2}{D_{eff}} \tag{5}$$

Even for a small R (0.001 cm), the time scale is over a day, which exceeds the duration during which most suspended particles remain in the water column. Hence the assumption of ***equilibrium partitioning is not really valid for resuspended particles***. Because much of the contamination on resuspended particles will not have time to de-sorb, assuming equilibrium partitioning overestimates the PCB flux from bottom sediments due to resuspension. Of course, during calibration, this could have been compensated for, in part, by assuming less sediment resuspension. **(1C)**

►*Biomixing and Bioavailable Depth*

Originally, the model used a bioavailable depth of 15 cm over which the biomixing coefficient (bio-diffusivity) was $D_b=10^{-9}$ m^2/s, or approximately 1 cm^2/d. I commented that this seemed quite large, and EPA agreed. They have changed their formulation to utilize a subduction velocity, V_s, which they get from a literature review of the rates of sediment mass reworking per organism per time, and site-specific data on organism abundances. Their chosen velocities are approximately $V_s \sim 10^{-9}$ m/s over the top 4-7 cm, and $\sim 10^{-10}$ m/s over the next 6-8 cm. In the absence of direct measurements of mixing, this may be the best approach, but it is noted from figures presented at the May 10 Document Review Meeting (DRM) that there is tremendous variability in organism reworking rates, suggesting much uncertainty.

Assuming a vertical distance of 5 cm between the top two sediment layers, their equivalent new values of bio-diffusivity are $\sim(10^{-10}$ to 10^{-9}m/s)(0.05m) $\sim 5\text{x}10^{-12}$ to $5\text{x}10^{-11}$ m^2/s or ~20 to 200 times smaller than previous. These are probably more reasonable, but it is difficult to assess whether or not they are right, because there is not much vertical variation in the existing sediment PCB concentrations. However, following remediation, the vertical concentration gradients could increase substantially, so this is an important process.

This also raises the issue of vertical resolution. If, following remediation, clean sediment is overlain by a thin layer of contaminated sediment, the numerical model will immediately mix the contaminant over the top layer (say 5 cm), whereas, with a bio-diffusivity of $5\text{x}10^{-7}$ cm^2/s, mixing will take $(5\text{cm})^2/5\text{x}10^{-7}$ cm^2/s or about 1.5 years to achieve this mixing. Thus, from a numerical modeling standpoint, ***the vertical grid size is too large.***

The sediment-water interface can rise or fall due to deposition and erosion, and EPA argues that this should not affect the rates of mixing. This may not be true since you could expect different rates of organism mixing in relatively fine grained sediments that have been recently deposited, versus older, more consolidated sediments that have been eroding. Indeed, Figure 13 of the DRM handouts on sediment mixing shows a strong increase in V_s with percent fines. However, to honor their assumption, they make V_s dependent only on the depth below the (moving) interface. But because the individual layer thicknesses are changing, the amount of vertical mixing will change. (A constant value of V_s will result in more mixing between thick layers having the same concentrations as thin layers.) Using a bio-diffusivity (dimensions of L^2/T) would take this effect into consideration since it is effectively V_s times the mixing length. **(3D)**

►***Diffusive exchange***

The calibrated mass exchange coefficient is K_f = 1.5 cm/d, which actually seems to be on the small side, since it incorporates a number of processes in addition to strictly pore-water diffusion. The coefficient for pore-water diffusion by itself should reflect the rate at which bioturbation brings PCB-sorbed sediment to the interface, the rate at which PCBs are desorbed to the porewaters, and the rate at which diffusion transports the dissolved PCBs into the overlying water column. These processes work like resistors in series. Chen (1993) showed that the sediment-water flux can be expressed as

$$J = \frac{C_{sL}/K_p}{\dfrac{\delta}{D_m} + \dfrac{2.2R}{(1-\phi)[(D_b + D')D_m \rho_s K_p]^{1/2}} + \dfrac{L}{(1-\phi)\rho_s K_p D_b}} \qquad (6)$$

where C_{sL} is the sorbed phase PCB concentration at the depth of the mixed layer L, δ is the water side boundary layer thickness, and D’ is the effective diffusivity of PCBs within the sediment (molecular diffusion as affected by tortuosity).

Jorgensen and des Marais (1990) suggest δ = 0.02 to 0.1 cm, and Shaw and Hanratty (1977) suggest

$$\delta = \frac{12 D_m^{0.3} \nu^{0.7}}{u^*} \qquad (7)$$

where ν is the kinematic viscosity of water ($\sim 10^{-2}$ cm^2/s) and u* is the friction velocity. Using u* = 0.25 cm/s gives $\delta \sim 0.05$ cm. Assuming L = 5 cm, $K_p = 10^5$ cm^3/g, $D_m = 0.5\text{x}10^{-5}$ cm^2/s, D' = 0.3x10^{-5} cm^2/s, $D_b = 5\text{x}10^{-7}$ cm^2/s (EPA's surface value), $\rho_s = 2.5$ g/cm^3, $\phi = 0.6$, and R = 0.01 cm, gives values for the three terms in the denominator of Eqn 6 of roughly 10^4, 3x10^1 and 10^2 s/cm respectively. This suggests that the flux is indeed water-side controlled (i.e, biomixing supplies contaminant to the interface sufficiently fast, and the contaminant desorbs sufficiently fast, that diffusion on the water side limits the transport) and that the last two terms can be ignored. Even if D_b were smaller by an order of magnitude ($D_b = 5\text{x}10^{-8}$ cm^2/s), the three terms in the denominator would be 10^4, 3x10^1 and 10^3 s/cm, leading to similar conclusions, albeit by a smaller margin. The second term represents the "resistance" due to desorption, and the fact that it is small suggests that equilibrium partitioning can indeed be assumed in computing the flux. The reciprocal of δ/D_m is K_f which, with the above numbers, is ~ 10^{-4} cm/s (8.6 cm/d). This is in the range of the values computed directly from EPA's flux analysis (e.g., Figure B.4-30 of the MCR, which includes values between 0.8 and 250 cm/d, with the majority between 3 and 10 cm/d). However, it is significantly above the value of 1.5 cm/d identified in the Phase 1 model calibration. One possible reason for the calibrated value being significantly lower is that the PCB concentrations used for the upstream model boundary conditions are generally higher than the data, at low flow, which may cause the calibration to underestimate the sediment-water exchange flux downstream.

Although the flux analysis indicated significant temporal variability in K_f that seems like it is correlated with stream flow rate, EPA's values of K_f based on complete model calibration appeared to be independent of flow rate. There seemingly should be some be some dependence, since increased flow would increase stream turbulence, decreasing δ and increasing K_f. As EPA acknowledges, as flow rate increases, it dilutes the water column concentration of PCBs, making it difficult to test for flow dependent fluxes. But observations do show that the model over-predicts water column PCB concentrations during low flow (when diffusive fluxes would dominate) which could reflect, at least in part, the lack of flow-dependence.

The time scale for natural recovery is the mass inventory per unit area, $C_s(1-\phi)\rho_s L$, divided by the flux, $J = C_{sL}K_f/K_p$, or $\tau = (1-\phi)\rho_s L K_p/K_f$. For $K_f = 10^{-4}$ cm/s (my value), and using other parameter values from above, the time scale is 160 years. Thus diffusive exchange may not be very important for contaminated sediments that are buried at or below the assumed level of bio-mixing (i.e., most existing sediments). However, for recently transported sediments with smaller L (e.g., following remediation), the time scale could be much smaller, especially if K_f were even larger. For example, if L were only 1 and K_f were simply twice the above value (2x10^{-4} cm/s), the time scale would be reduced to 16 years. However, the model would not be able to resolve the resulting sharp gradients with the current, relatively coarse, vertical grid scheme. **(1B)**

►***Erosion***

The erosion formulation and parameters come from analysis of SedFlume data. W. Lick argues that the exponent n (denoting dependency of erosion on shear stress) should be significantly greater than the chosen value, which would produce proportionally more erosion during high flow conditions. I agree with him, but am also leery of the fact that shear stresses in the model are computed based on a grid width that is essentially equal to the river width. Thus the computed hydrodynamics will yield cross-sectional average velocities, which ignores regions within a cross-section with relatively high velocity and erosion (refer to later discussion under Question 7). Unlike conditions in SedFlow, the real river is non-uniform. Thus the formulation of erosion can not be divorced from the question of grid resolution. Indeed, it would seem incorrect to simply import an erosion calculation from SedFlume. **(1A)**

►***Spatial Variability***

Much has been 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. Whether or not this is a model failure, *per se*, or simply unresolved variability in model input and output (sediment bed concentrations), depends on what has caused the variability. I believe the variability was caused mainly by the stochastic method in which the PCBs were introduced in the first place. In such case, 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 (and some non-linearity is inevitable, given the averaging associated with the coarse grid). 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 (at least for natural attenuation).

On the other hand, if the variability is due to active sediment transport processes that are sorting the sediments and their contaminants, then the failure to pick this up could be a significant model deficiency. For example, natural attenuation could conceivably increase local PCB concentrations. The available time series data of PCB concentrations within surficial sediments suggest a decrease in concentration (indeed more so than is being modeled), so I don't believe this is a significant process. **(2B)**

►***Boundary Loads***

The trend of the PCB concentration versus flow data for the E. Branch, shown in Figure 4.1-2 seems strange (why the sudden change at 550 cfs?). I guess this is simply what the data suggest. Also the model fit exceeds the data at low flow (10 cfs) by a factor of 2-3 for both dissolved and particulate PCBs. This excess upstream load may have led to an underestimation of calibrated fluxes from other sources downstream. Given the prevalence of low flow periods, this is important. **(6)**

►***Initial conditions for sediment PCB concentrations***

I agree with GE that hindcasting predicted trends is not a robust way to establish initial conditions. By playing the tape backwards, then forwards, they will get the same conditions they started with. I am not sure what the alternative is, but the lack of independence should be acknowledged. **(7)**

►***Summary for Question 1***

I believe that several processes could be better represented. But I recognize that several of these processes have been calibrated, so that ***if one parameter were changed, others would have to be changed*** as well. If there is time, I would like to see the model recalibrated using more appropriate values, within physical/chemical/biological constraints. (I do not think an independent validation is necessary.) In the absence of re-calibration, the model users must exercise considerable judgment in the interpretation of model output. **(13)**

Question 2 Are the comparisons of the model predictions with data sufficient to evaluate the capability of the model on the spatial and temporal scales of the final calibration and validation.

Question 3 Is there evidence of bias in the models, as indicated by the distribution of residuals of model/data comparisons.

These questions are clearly related, so their answers are combined. The “upstream” and “downstream” models are discussed at the end.

►***Flows and Velocities***

Flows were simulated for the 10.5 year Phase 2 calibration period (Figures 4.2-4 through 4.2-14). Unfortunately there is not much data to compare with. The most data are for early 1999 when the model consistently under-predicts flow. (Yet it seems to do well in the later validation periods, based on the pressure transducer data.) Similar agreement with stage suggests that average velocities should be pretty good. (Figure 4.2-26 shows reasonable agreement between measured and predicted velocities, but the predictions are generally too low; EPA argues that discrepancies are due in part to coarse grid resolution.) During storms there was reasonable agreement, with moderate overprediction of flow at low flow. Agreement during storms shows times of over- and under-prediction, but they seem OK on average. The model seems to do reasonably well simulating the extent of overbank flow during August 1990 (Figure 4.2-25). Model statistics seem generally within ranges specified by the QAPP and there are few indications of bias.

In the validation period, flows and velocities look similar to or better than the calibration period. There are some errors in the timing of flows, but these are not important. No major bias is seen. **(3B)**

►***TSS***

The model uses flux analysis for upstream TSS when there is no data and often this analysis overestimates TSS by a factor of 2-4 (Figures 4.2-31 and 6.2-12). There is some underestimate of TSS at high flow.

The model tends to overpredict TSS during the calibration period (Figure 4.2-39), especially at lower TSS concentrations. This is possibly because of the overprediction of upstream loading. The model seems to do better with the storm events, which EPA claims are more important. While bias in TSS at low flow may be insignificant in terms of the sediment budget, it does affect the water column PCB concentrations, and hence the relative uptake of PCBs from the water column and the sediment in the Food Chain Model. In general it would have been nice to have had more data on TSS.

The statistical summary (equivalence plot, Figure 4.2-61) indicates generally good agreement, but the model shows less variability than the data; i.e., the model overpredicts low concentrations and underpredicts high concentrations. There appears to be no significant bias with flow rate. Sedimentation rates agree reasonably well with measurements using Cs-137.

Bedload concentration is computed as bed load mass rate divided by river flow (pg. 4-56). Given that the bedload is traveling slower than the average stream velocity, this would seemingly give too low concentrations, but I suspect velocities are more important than concentrations. **(6)**

►***PCBs***

The spatial plots at low Q (Figure 4.2-69) and moderate Q (Figure 4.2-70) show that upstream concentrations of both TSS and PCB are too low (a consequence of the BC) and that there is more local variability (indicated by shading) than variability between reaches. There is no way to verify the local variability. It is hard to say whether the overall spatial trend is correct or not, because of variability in the data. The temporal trends look OK.

The temporal changes in sediment PCB (top 6 inches) show that the model does not reproduce the significant decline in Woods Pond over time (Figure 6.2-50). This is one of GE's main points and suggests that the model might also underpredict the response to remedial actions, including natural attenuation. ***I see this as a major concern***.

The model misses the decline in PCBs across Woods Pond that is both observed and intuitively suspected. This would mean the model overpredicts loading downstream (downstream model).

Like TSS, the equivalence plots for PCBs (Figure 4.2-85) shows less variability than the data. However, there is little indication of (overall) bias.

And also like TSS, it would also be nice to have had more data on PCBs. Again, there is better model-data agreement at relatively high (presumably more biologically relevant) concentrations. As with other variables, the model shows less variability than the data (i.e., it underpredicts the highs and overpredicts the lows).

Except for the transducers which record flow (through a stage discharge relationship), better data seem to be available for the calibration (especially the original calibration) period, than for the validation period. However, I am less concerned with traditional model validation (the model's ability to predict absolute concentrations under different hydrologic conditions) than I am with the model's ability to distinguish among different remedial alternatives that would presumably be evaluated under the same hydrologic conditions. See comments under Question 6. **(2A)**

►***Upstream Model***
The model was extended two miles upstream to include the 0.5 and 1.5 mile remediation areas. Upstream Q and TSS were obtained from rating curves, while upstream tPCB = 0.019 microgram/L from boundary measurement. Comparisons were attempted against TSS and tPCB measured at Lyman St. in 2001-2004 (some of this is reported in DRM handouts and not in the MVR).

Predicted TSS varied from ~$3x10^{-4}$ to >10^2 mg/L while data were in the range of 3 to 30 mg/L (mostly 3-10). [I am assuming these are the correct units for TSS; the vertical axis on the handout distributed at the DRM says ng/L.] Can one really predict TSS as low as $3x10^{-4}$ mg/L? Some data are in the range of the model, but model predictions show much more variability? Where does this come from?

Predicted PCBs were $3x10^{-5}$ to 10^{-3} mg/L while data were in the range of ND up to $3x10^{-5}$. [Again, I am assuming these are the correct units; they are consistent with the 30 to 1000 ng/L indicated on the handout GE provided, though EPA's handout during the document review meeting indicates $3x10^{-5}$ to 10^{-3} ng/L.] In any case, the agreement is not very convincing, and ***it is hard to say that the upstream model is validated***.

The downstream concentrations of TSS and PCB from the upstream model are much lower than the corresponding upstream boundary values used for the main model. Given that output from the main model depends so heavily on its upstream boundary condition, more work should be done on the upstream model, or at least more interpretation of the results. There is scant discussion in the MVR. **(8)**

►***Downstream Model***
The model was extended 19 miles downstream to Rising Pond Dam (not described in MVR; only in DRM handouts). Upstream boundary conditions were taken from output of the main model at Woods Pond. Model-data comparisons for TSS and tPCB were made between 1990 and 2004 ~ 1mile downstream from Rising Pond Dam.

Predicted TSS varies widely as do the data, with the peaks corresponding with peaks in flow. They are in the same ballpark, but the variability is so great that you need a statistical comparison to determine how well they match on average.

Predicted PCB concentrations also vary substantially (~3×10^{-3} to ~3 microgram/L), with predicted peak concentrations corresponding with peak TSS and peak flow, and non-peak concentrations hovering generally around 0.03 micrograms/L. (But why are some up to 10 times lower?) Measurements show much less variability with an average of around 0.1 microgram/L, or about 3 times the predicted non-peak values. The data are too sparse to see if they respond to peaks in flow. Again a statistical comparison would be helpful to quantify model-data agreement. Both measurements and predictions seem to be 2-3 times lower than corresponding values at Woods Pond Outlet, which makes sense given the 20 mile separation.

GE's plot of simulated versus measured PCBs at Rising Pond Dam shows very little correlation between simulation and measurement ($r^2 = 0.01$). ***At this point it is hard to say the downstream model is validated***. **(9)**

Question 4 Have the sensitivities in the models to the parameterizations of the significant state and process variables been adequately characterized?

►Most sensitivities seem reasonable. As expected, flows show strong sensitivity to upstream boundary flows, TSS is sensitive to upstream boundary flow and TSS, and PCBs are sensitive to upstream boundary flow, TSS and PCB concentration.

At low flow PCB there is strong sensitivity to K_f and partitioning (at least at New Lenox Rd), which makes sense, since "diffusion" is the only important process, but this effect is largely diluted out in the 10.5 year simulation. At high flow, there is strong sensitivity to parameters dealing with particulate phase transport (settling, erosion).

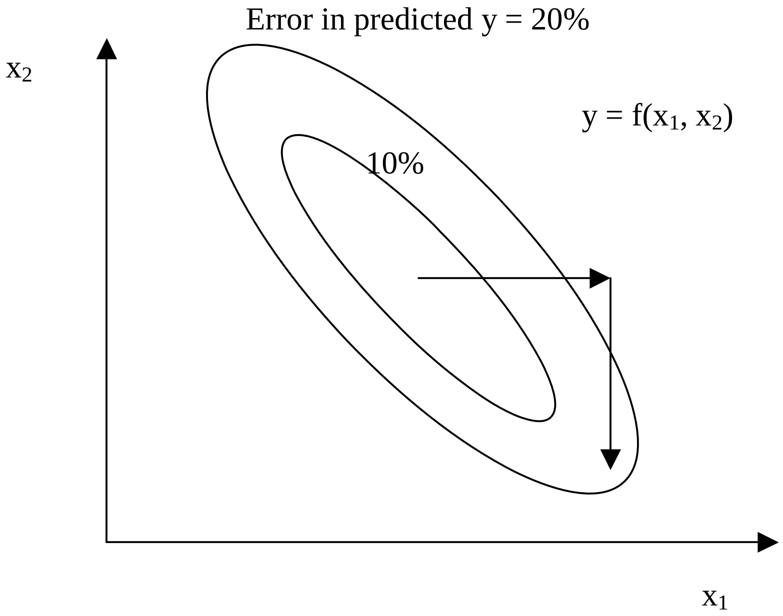
While the model shows substantial sensitivity to diffusion parameters during low flow at New Lenox Rd, GE points out that virtually no sensitivity is indicated at Woods Pond Footbridge. Since the flow must pass from NLR to WPF, this is illogical. Unless there is a mistake in plotting (and EPA suggested there was not), this could indicate a potential problem in model formulation. **(5A)**

Question 5 Are the uncertainties in the model output(s) acknowledged and described?

►In general, model uncertainty comes from two sources: imperfection in the basic model(s) (including unknown unknowns) and imperfection in the selection of the (known) model parameters. The former has been discussed to a certain degree under Question 1, which relates to processes, so the following relates mainly to the latter issue.

Although they use some sophisticated approaches, the modelers performed a rather traditional uncertainty analysis, varying known model parameters largely ***independently*** in a simulation of ***existing conditions***. If the model and data were both perfect, and the only uncertainty involved the values of model parameters, one could expect most of the parameters to be independent. (Some parameters that have a physical relationship, such as erosion parameters would still be correlated.) But the model has been calibrated to observed data, and the calibrated model parameters reflect both imperfection in the model (e.g., a parameter being used to cover more than one real process) and imperfection in the data (e.g., due to sampling or analytical error, insufficient sample resolution, etc). Thus, during uncertainty analysis, when one parameter is changed, other parameters should also change (i.e., they are not independent), as is illustrated in the following sketch. Here the dependent variable y is a function of two independent variables x_1 and x_2 and the contours are plots of some measure of error between measured and predicted values of y. If, as part of an uncertainty analysis, variable x_1 were changed (horizontal arrow), one would expect a correlated change in x_2 (vertical arrow) in order to minimize the error between measurements and predictions. Failure to take this correlation into account tends to overestimate the uncertainties in the model prediction. And, as suggested in the following paragraph, I suspect this issue becomes more extreme when the future conditions being predicted are substantially different from those under which the model was calibrated.

Apart from the question of parameter independence, there is the question of whether there are certain variables to which the model is more or less sensitive as a function of the remediation scenario. For example, following remediation there may be thin patches of clean/dirty sediment overlaying thicker regions of dirty/clean sediment, making calculations more sensitive to diffusion and partitioning than they were in the base case scenario. If this is the case, we may get a distorted impression of the effectiveness of remediation.



Based on sensitivity analysis, EPA tested the uncertainty of model outputs (e.g., PCB concentrations in water, bed and fish) to uncertainties in model input parameters. As GE points out, the model uncertainty is substantially greater than the uncertainty in the model (e.g., the s.e.m.) If the goal were simply to predict the uncertainty in present conditions, the uncertainty in the individual parameters could be scaled back until the uncertainties in the model and data matched. But it is also important to see the ***sensitivity of predicted model changes due to possible remedial alternatives*** to these same model input parameters. This is not difficult and should be performed either before or after the model is transferred to GE.

Finally, one use of the model is simply to compute relative fluxes, as EPA as demonstrated recently. Remedial decisions could be made on the basis of how much less PCB would be transported, rather than on (or in addition to) the basis of quantitative predictions of concentrations. Thus it would also be nice to see the ***sensitivity of predicted fluxes to the same model parameters***.

In summary, because of the lack of parameter correlation in the sensitivity study, and the fact that predicted uncertainty exceeds the uncertainty in the existing data, I believe the predicted uncertainty, when applied to future remedial conditions, will be too large. **(5B)**

Question 6 Upon review of the model projections of changes in PCB concentrations in environmental media in the example scenarios, are such projections reasonable, using your technical judgment, and are they plausible given the patterns observed in the data?

►I appreciate the inclusion of the two example simulations that shed some light on how the model might perform when simulating remediation options. This is clearly a start. There would seem to be two issues to consider. The first is how well the remediation measures themselves will work. For instance, will a cap erode? The second is what will happen downstream (in space and time) after remediation? For example, will dirty sediments cover up capped or dredged sediments? The example simulations don't consider the first question (though it is very important), but partially address the second. However, several considerations come to mind. First, the example simulations were begun with "cold start" initial sediment conditions, which seems to explain, at least in part, why sediment concentrations increased in some areas downstream from simulated remediation (i.e., places where there was no PCB source). Better that the model be spun up to initial sediment conditions that reflect an equilibrium with the modeled hydrology and bathymetry. Also, the model seems to ignore any particulate PCBs that might have wanted to deposit in the channel of Reach 5A. (What happens to these in the model?) Ultimately, the model could be generalized to look at "different colored" PCBs emanating from different source locations and to see where they end up, with implications as to which would be the best sediments to remediate. This might require subscripting certain variables in the code so the calculations could be run in parallel. EPA has started to do some of this, but more would be instructive. **(10)**

►As I expressed at the DRM, I am concerned with the absence of validation data appropriate to remedial measures. Dredging would result in horizontal gradients in sediment concentration over 10s to 100s of meters, while capping would result in vertical gradients in sediment concentrations over a few cm. The current sediment concentration distributions do not show variability over these scales, so it is difficult to tell whether or not the model will be able to successfully predict the effects of such remedial measures. **(12)**

►Figure 4.2-53 shows areas of net deposition indicating ~ 42% of net deposition is to floodplain, and ~ 10% is to Woods Pond. The remaining ~48% is to the sediment bed upstream of Woods Pond. It would be nice to keep track of the deposited sediments. My guess is that those deposited to the flood plain and to Woods Pond stay put, or are slowly released, while those in the sediment bed move on. Can this be quantified? **(10)**

Question 7 Is the final model framework, as calibrated and validated, adequate to achieve the goal of the modeling study to simulate future conditions 1) in the absence of remediation, and 2) for use in evaluating the effectiveness of remedial measures?

Comments focus on the related questions of grid resolution and computational feasibility. Recommendations follow at the end.

►Like most other review panelists, I continue to be ***concerned about the lack of lateral resolution in the computational grid***. Having only about one grid cell over the river width means that important processes must be parameterized. Because the existence and magnitude of erosion and deposition depend non-linearly on stream velocity, a trapezoidal section with depth varying linearly between 1 and 3 meters at the two banks will look to the model like a section with uniform depth of 2 m, yet in practice the former may have regions of strong erosion and deposition (as in re-surveyed cross-section XS061 shown in the DRM handouts), whereas the latter might be marginal one way or the other. Even if the model gets the net erosion/deposition correct (doubtful due to the non-linearity), the failure to capture gross erosion/deposition is problematic, both for evaluating remedial measures (such as capping) and natural attenuation. Assume an entire channel cross-section is capped. If half of the channel is erosional and the other half is depositional, the first half will require added protection to keep the cap intact, while the second half will get covered with contaminated sediments from upstream. Neither process would be forecast with a model that predicts marginal net erosion/deposition. As for natural attenuation, contaminated sediments from an upstream area with gross erosion can contaminate (or re-contaminate) downstream areas, but the model would not predict this. **(3A)**

►The issue of model resolution conflicts directly with the issue of computational time. Clearly, with the current coupled in-channel/overbank modeling system, one cannot afford the extra computational cost of reducing Δy, and still be able to afford multiple simulations of multiple scenarios each over multiple decades. I share GE's concern that the existing model is already too cumbersome to be used effectively to study corrective measures. Indeed, this is probably the biggest model issue: as Yogi Berra might say, "If you can't use the model, you can't use the model." The following are some thoughts on the related issues of grid resolution and computational cost. **(4)**

►***Grid resolution***

Too begin with, it must be recognized that the lateral grid size is a model parameter as well as a numerical parameter. Ideally, when a set of equations are solved using a discretized numerical model, one likes to reduce the grid size until results converge. ***But, we are no where near that here.*** Hence if the model grid size were to be reduced, this would affect other model processes (notably erosion, which depends heavily on velocity, which would vary significantly across the river width). Thus ***if the model grid were to be reduced, the entire calibration would need to be redone***.

It is not clear how much resolution would be needed to properly compute transport. Because of this uncertainty, and the time and effort involved, it may be too late in the game to make major changes to the model grid. However, at a minimum, some sensitivity tests could be conducted using a finer grid over a small portion of the (in-channel) domain and run over a short duration of the 26 year simulation period. The output from the more highly resolved model could be compared with that from the coarser model and then parameterized. For example, it might be determined that the net erosional flux with a resolved grid is X times that computed with the coarse grid, in which case predicted erosional fluxes with the coarse grid would be multiplied by X. **(3A)**

►***Computational costs***

A number of options have been mentioned by the Peer Review Panel and by GE to improve model efficiency. These included splitting the calculation of hydrodynamics, sediment transport and PCB fate; running synthetic hydrological sequences; and employing alternate grid schemes. EPA did implement some efficiency measures (dynamic and split time steps, by-passing sections of the model that change slowly and running simulations of selected portions of the calibration period), but seemed to have dismissed the bigger ticket items. I think some of these have to be revisited.

There is no reason, in principle, why the model calculation of hydrodynamics, sediment transport and PCB transport/fate could not be de-coupled. In particular, the hydrodynamics could be run first. These results could be stored and used to transport sediment and PCBs. Since the hydrodynamics would only need to be computed once (hydrology does not depend significantly on scenario), considerable saving could be obtained. The devil is in the details of course, and this would probably take a few months of time. But since EPA is a model developer, and such a decoupled model would have future applications, EPA may want to invest in such a project, possibly calling in the model developer (J. Hamrick).

Further savings could be obtained ***using synthetic hydrologic flows***. Based on the relatively large data base, a suitable discrete distribution could be generated that includes representative flows of different magnitudes and recurrence intervals. The model (with or without the hydrodynamics disaggregated from transport) could then be run for a synthetic year (or a short period of years large enough to include the largest/least frequent flow of interest), and the changes per year documented. Perhaps, without rerunning the model, ***the long term effects could be computed by extrapolation of the one year results***.

Also, it still seems to me that in channel and over bank calculations could be de-coupled, saving additional computational time, and/or allowing more detail during the vast majority of time when the flow is within the banks and essentially one-dimensional. EPA claims there are coupling issues such as loss of momentum conservation, which may be true. But exact coupling is not critical. We are only expecting the model to predict gross trends in PCB concentrations (over time and the longitudinal direction), so a temporary mis-match might be acceptable. In general, I would support more a somewhat more approximate model (e.g., with parameterizations) that could be more efficiently used.

In this regard, I like M. Garcia's idea of a flood plain number to quantify deposition of PCBs within portions of the floodplain. PCBs appear to be on a conveyor belt whereby they are resuspended from within the channel at high flow, partly deposited on the banks, then gradually eroded from the bank back into the river, until some of them are deposited in Woods Pond. These processes are all in the model, but it might be nice to document the life history of a "numerically marked" cohort of sediment, and to test the sensitivity of the ultimate fate of this cohort to processes such as the bank erosion rate, frequency of over bank flows, etc. By decoupling the sediment transport and fate, this could lead to a model simplification, whereby the over bank processes (deposition and bank erosion) are parameterized as simple first order sinks and sources whose rates are determined by calibration to detailed simulation over the calibration period. These rates could vary with flow conditions, but would be independent of PCB concentration, allowing simple calculations during the Corrective Measures Study.

In summary, it is a shame that the longitudinal and transverse grid sizes (Δx and Δy) must be similar since the former need not be nearly so small (fish average their exposure over large distances) and the latter should be smaller. In my previous comments I suggested eliminating Δx as a dependent variable (substituting, instead, grain size by reach), but I recognize this was a radical idea that might take too much time to implement.

Finally, while the above discussion refers to lateral resolution (and its relationship to computational cost), I am also concerned that the vertical resolution within the sediment bed that may be insufficient to resolve near bed gradients that might result from remedial efforts (see Question 1). Clearly, calculations in the sediment bed can use a different (much longer) time step than those in the water column. It is not clear how costly additional vertical resolution in the sediment bed would be, but I think it should be considered. **(4)**

Recommendations

►Considering the difficulty of the modeling task, the current state-of-the-art, the amount of effort that has already gone into the modeling, and the difficulty in arriving at a consensus (when our review panel is not allowed to), I would say that ***the current work is "acceptable"***. However, in view of the many reservations that many of us have, the model users (GE) will have to exercise considerable professional judgment in their use of the model, and EPA should grant them this judgment. **(12)** ►And depending on how much time is available, ***I would recommend some modest additional effort***, which I organize into three categories: easy, intermediate and difficult. I believe that GE should be involved with this as much as possible (rather than simply be given a black box), because it is their model judgment that will be relied upon, and this judgment can only come with experience using the model.

Easy

These are things that should be easy to accomplish within a framework of a few weeks to a couple of months. They are probably also the most important.

1. Develop a strategy to improve model efficiency so that remediation scenarios can be evaluated efficiently. I suspect that the best approach here is to make use of synthetic hydrological sequences that are repeated.
2. Perform additional uncertainty analysis that looks at the uncertainty of future remediation scenarios (as opposed to simply past natural attenuation), both in terms of absolute concentrations and in terms of PCB fluxes from different sources.
3. Improve the "upstream" and "downstream" models through further calibration (GE indicates that additional data are available for this purpose).

Intermediate

The time frame here might be several months to half-a-year.

4. Develop a more highly resolved grid within at least a portion of the channel and use the grid in a short term simulation to understand the sensitivity of model predictions (again, actual concentrations as well as fluxes) to grid resolution. Information parameterized from this sensitivity test could be used to adjust the coarse grid output, as suggested above.
5. Decouple the model of the channel and the floodplain (to improve model efficiency).
6. Decouple the hydrodynamics from the sediment transport and PCB fate within EFDC (to improve model efficiency).
7. Re-calibrate the model based on suggestions made by reviewers under Question 1, and in consideration of the several instances of bias noted under Question 3. (I don't believe a separate validation is required.)

Difficult

This step would lead to the greatest model accuracy and robustness, but might take a year or so to accomplish.

8. Restructure the grid to allow more lateral resolution. This would have to include computational efficiency measures so that extended model runs are feasible, as well as a complete model re-calibration. **(13)**

References

Chen, H.-W. (1993). Fluxes of organic pollutants from the sediments of Boston Harbor. S.M. Thesis, Dept of Civil and Environmental Engineering, MIT, Cambridge, MA.

Jorgenson, B.B. and D.J. des Marais (1990). The diffusive boundary layer of sediments" oxygen microgradients over a microbial mat. *Limnol. Oceanogr*. 35(6): 1343-1355.

Shaw, D.A. and T.J. Hanratty. (1977). Influence of Schmidt number on the fluctuation of turbulent mass transfer to a wall. *A.I.Ch.E.J*. 23, 28 pp.

Wu, S.C. and P.M. Gschwend. (1986). Sorption kinetics of hydrophobic organic compounds to natural sediments and soils. *Environ. Sci. Technol*. 20: 717-725.

Wu, S.C. and P.M. Gschwend (1988). Numerical modeling of sorption kinetics of organic compounds to soil and sediment particles. *Water Resources Research* 24(8): 1373-1383.

APPENDIX A.2

COMMENTS OF W. FRANK BOHLEN

Model Validation:

Modeling Study of PCB Contamination In The Housatonic River

A Peer Review

By

W. Frank Bohlen
University of Connecticut
Department of Marine Sciences
Groton, Connecticut 06340

Member, Peer Review Panel

July 20, 2006

Introduction

Under agreements developed between the General Electric Company and the U.S. Environmental Protection Agency future designs of remedial activities intended to reduce exposure to PCB contaminated sediments within the Housatonic River will make use of a predictive numerical fate and transport model. This model, under development since 2000, is intended to provide quantitative measures of sediment and PCB transport and associated uptake by selected biota over a variety of spatial and temporal scales. Initially, the model has been applied to a region extending downstream from the confluence of the East and West Branches of the River (approximately 2 miles downstream of the GE facility in Pittsfield, Ma.) To Woods Pond Dam, a distance of approximately 10.7 miles. This Primary Study Area (PSA) is believed to contain 90% of the mass of PCBs present in the River. The PSA is a morphologically complex area which in combination with regional hydrology and placed control structures (i.e. dams) establishes a multi-faceted transport regime. The complexity of this system, representing a particular challenge, has lead to the development of a model consisting of three primary elements, a watershed model (HSPF), a hydrodynamic/sediment-contaminant transport model (EFDC) and a bioaccumulation model (FCM). As presently configured, the models are linked but not interactive.

Model development proceeded first through the conceptual phase (~2000-2002) and then through calibration (2002-2005). Over the past year emphasis has shifted to validation. This included a revised calibration effort in which the initial 14 month calibration was extended to a 10.5 year period (Jan.1990-June 2000), a number of model changes to better treat key processes associated with the complexity of the transport regime, and then validation over a 26 year period 1979-2004. Coincident with validation the model domain was extended in space to include a region upstream to Newell Street in Pittsfield, Ma (a distance of approximately 1.5 miles up the East Branch of the River from the Confluence) and a region downstream from Woods Pond to Rising Pond in Great Barrington, Ma. , a distance of 19 miles.

The results of the revised calibration and independent validation were released in a two volume report in March, 2006 (Weston, 2006) for public comment and review by a Peer Review Panel. The following provides a summary of my review of this report, supplementary materials provided by EPA associated with the Document Overview Meeting held in May, 2006, and review comments submitted by General Electric Contractors and concerned citizens.

General Comments

▶This has been and remains an ambitious project. In hindsight it may have been overly ambitious to expect a single model formulation to efficiently and accurately predict future PCB concentrations throughout a morphologically complex region over extended periods of time given what was known about historical distributions and the range of processes governing transport and fate. Some of this seems to have been recognized by EPA in their recent refinement of the modeling goals placing primary emphasis on the need for the model to be able to establish the relative performance of selected remedial alternatives rather than on its ability to yield certain prediction of absolute PCB concentrations (EPA, 2006). While this refinement is advisable and will be taken into consideration in model evaluations it must be remembered that it does not relieve the modeler's responsibility to develop an efficient, stable and quantitatively accurate model. The evaluations to be conducted during the upcoming Corrective Measures Study (CMS) by GE will be considering the benefits of remedial alternatives over extended periods of time, often 40 years or more, i.e. times well in excess of the validation period. The utility of the relative comparisons over these extended periods will ultimately depend on the degree to which the model provides accurate simulation of all of the governing physical chemical and biological factors affecting transport and fate and the adequacy of the computational numerical schemes. Fundamentally, these are the same factors to be considered if the model was to be used for absolute predictions. The sufficiency to evaluate these characteristics depends in large part on our understanding of the PCB transport system structure and dynamics within the Housatonic River. **(12)**

▶On several occasions discussions with the model group and within the Peer Review Panel made it clear that this required understanding of the variety of transport processes affecting PCB transport in the PSA was less than perfect. Recall the discussions of floodplain dynamics, an area known to represent a significant sink and possible longterm source of PCBs, the continuing debate over sidebank transport,the specification of boundary conditions, and the proper structuring of the sediment bed within the model. Each of these items represent an essential element of the transport model and most will come up again individually in the following review of the model validation effort. Viewed collectively this variety of unknowns makes clear that what EPA and GE are dealing with is a research project rather than simple application of an accepted formulation. With this fact in mind the future role of the model should change to include guidance of monitoring efforts. The need for these additional data should be clear from a scientific standpoint. Their availability would also serve to increase stakeholder confidence in model results.

Monitoring to date has placed primary emphasis on PCB distributions within the sediment column with relatively limited sampling of the water column TSS and PCB concentrations. The latter have placed primary emphasis on flow/TSS relationships during high flow events with sampling at a number of selected transects along the main stem of the river. This monitoring has been supplemented by some few field observations of sidebank erosion and laboratory estimates of bed erodability using SEDFLUME. With the exception of the sidebank observations the majority of the field observations have not been directed at specific processes.

I'd recommend that consideration be given to the extension of these past monitoring efforts to include, for example, the placement of instrument arrays at the Confluence and at the Woods Pond Dam sufficient to provide long term, high frequency (e.g. 3-4 samples/hr), time series observations of water temperature and TSS at the mid-point of the low flow water column. These measurements would be supplemented by monthly sampling of concurrent PCB concentrations. All instrument observations could be telemetered to a central station permitting conditional sampling as unusual flow/transport conditions occur. In addition to the upstream and downstream stations in the PSA consideration should be given to the placement of one or more instrument arrays at selected sites adjoining the flood plain. Again these relatively high frequency data should be supplemented by lower frequency drawn water sampling of concurrent PCB concentrations. This latter sampling might occur on a monthly basis and aperiodically during particular rainfall/runoff events. This combination is intended to significantly increase our understanding of flood plain transport processes and their temporal (including seasonal) variability. These observations would also take allow us to take full advantage of the ongoing remedial efforts by providing quantitative data detailing effect at a number of locations. Such data would seemingly be of value in future remedial planning.

The above observations should be supplemented by a variety of other process studies such as the continuing survey of selected portions of the side banks and sequential bathymetric survey of river channel transects or detailed cross-channel velocity/flux measurements to establish the adequacy of the model grid. Model results would be used to specify siting as well as the need for continuing observations. This close coupling between models and monitoring would be of clear benefit to the long terms goals of this effort.

Beyond these technical issues, the reports provided this reviewer remain exceedingly difficult to read. I understand all of the reasons why but cannot believe that the project would not benefit from a clearer and more concise document. In this validation report there is entirely too much use of “reasonable agreement” and the like with often insufficient demonstration. The Executive Summary is too general and does little to build confidence in this modeling effort and its subsequent application. Questions regarding many of the key points of the model require going back to previous documents that were themselves obscure and it’s often difficult to figure out just what is being presented in the report figures (e.g. was that streamflow instantaneous, daily average or monthly average ? Are the TSS values a vertical average? Over what period of time? Is it legitimate to compare longer term model results to shorter term data?). Too many discussions end prematurely before any attempt is made to explain observed differences or discrepancies (see pg. 4-90/91 Vol.1 discussion of Event 10. Why the underestimation?) Many of these questions might be resolved by a search of our voluminous file but who but the most dedicated would be expected to do it?

I’d recommend, now that the major components of this exercise are in place, that a technical writer be charged with the preparation of a single document describing the model and the resulting runs written for the stakeholder community. This document would include all major features of the model and results with key supporting figures, references and an index. I’d consider this a high priority. **(13)**

Moving now to the specific questions posed to the Peer Review Panel:

1. Considering the changes implemented in the Phase 2 Calibration, does the model as calibrated and validated, based on your technical judgement, reasonably account for the relevant processes affecting PCB fate, transport, and bioaccumulation in the Housatonic River to a degree consistent with achieving the goal of the modeling study?

▶The global list of processes and governing factors incorporated in the linked three (3) model system is comprehensive and includes all those necessary for a detailed evaluation of PCB transport, fate and bioaccumulation (see Table 2-1 pp.2-3 Vol.1). In addition, the model, as structured provides an excellent framework for the systematic evaluation of each of the factors governing PCB transport and its ultimate bioavailability in the Housatonic River. This framework directly complements quantitative study of transport and subsequent investigation and ranking of remedial alternatives.

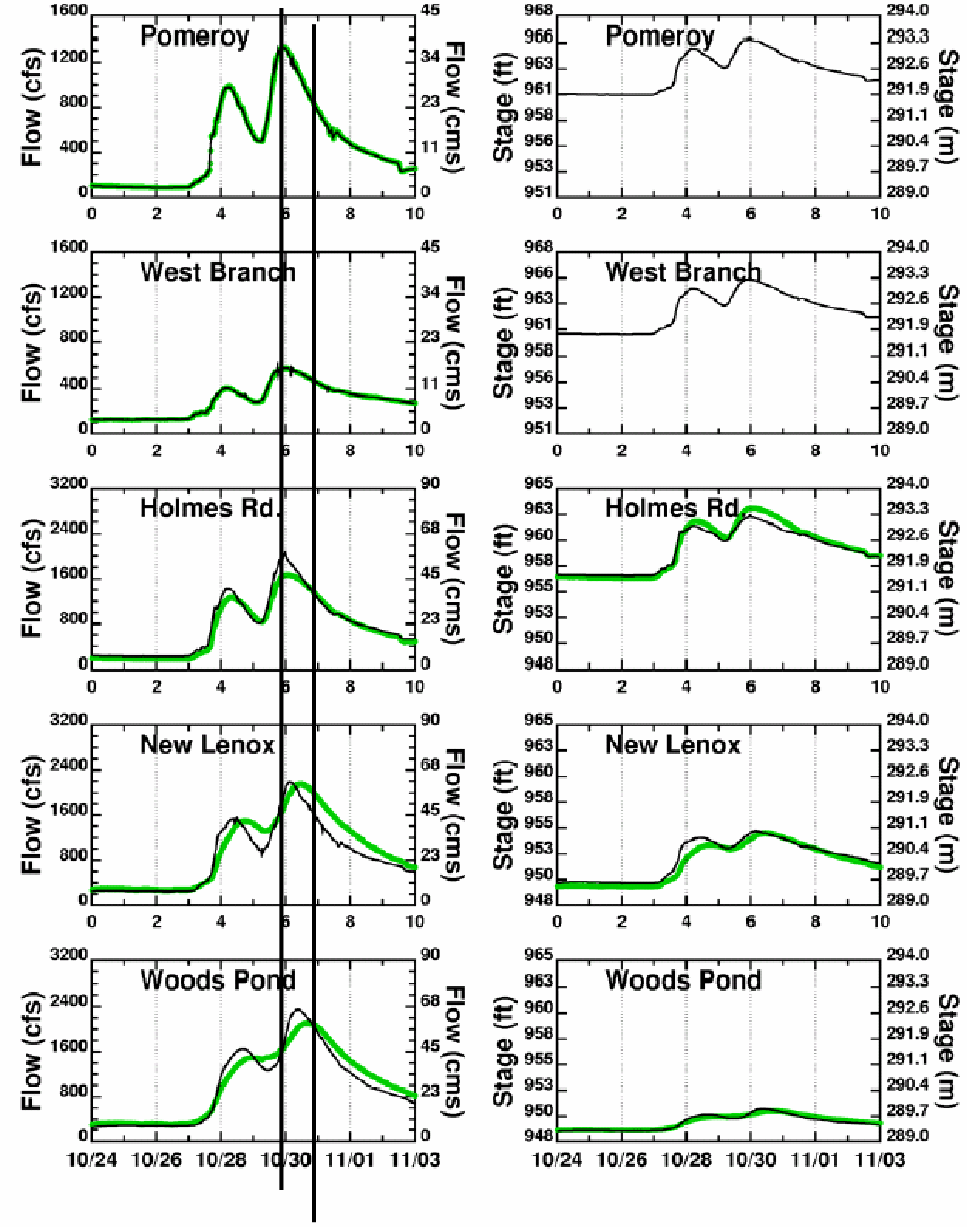
Despite the completeness of the global list, however, realization of the full potential of the models is governed by the extent to which model formulation provides accurate process simulation. Review indicates that model utility would benefit from improvements/modifications in a number of areas. **(1)**

▶In general, comparisons with observed discharge indicates that the watershed model (HSPF) provides accurate simulation of the factors governing stream flow volumes to and through the PSA. These comparisons also suggest that the model is able to reproduce the timing of flow events. This matter of timing is an important factor if the data are to be used to calculate velocity and ultimately boundary shear, as they are in this study and will be in the upcoming

Corrective Measures Study. The actual timing of stage/discharge at each section of the study area in large part determines the magnitude of the horizontal pressure gradient which affects speeds, turbulence intensity and boundary shear. These are the principal factors governing sediment/PCB transport both in the water column and across the sediment-water interface.

Given the importance of timing it is disappointing that the report provides so little detailed information on this factor and its effects. There is abundant reference to the model's ability to accurately reproduce event timing but these statements appear to be referring to timing in the most general sense. i.e. a precipitation event induces an increase in streamflow over a time similar to that observed. These references to "event timing" leave the question of the adequacy of the model simulation of timing with respect to flow velocities unanswered. Examination of several of the figures (see Fig. 6.2-5 (attached) e.g.) shows substantial differences between measured and modeled actual flow and stage timing at several stations along the PSA. In this figure, it would appear that the modeled speeds produced by the associated pressure gradients should be less than the measured and that the turbulence induced by adverse pressure gradients would also be reduced. Actual estimation is, of course, complicated by the morphology of the region with tributary inflows and/or backwater flows and storage complicating the simple stage discharge relationship. This factor may be the reason for example that the simulated flows at Holmes Road are higher than the measured despite a higher measured stage relative to that simulated. **(13)**

▶This matter of flow phase and velocity and associated effects on transport could be better evaluated if the report had provided a more complete discussion of measured vs. modeled velocities when presenting the results of EFDC hydrodynamic calculations (see pp.6-31 e.g.). Although the model uses HSPF generated flows at the upstream boundary, the resulting simulations are to some extent independent of the HSPF generated flow/stage through PSA and as a result may be less sensitive to this matter of stage timing. Speeds were measured using both electromagnetic and acoustic doppler currents meters for short periods of time at several locations within the PSA (see Figs 4.2-26 and 6.2-8 (attached)). These comparisons, while taken to be generally acceptable, indicate to me that there is very likely a substantial difference in measured versus modeled sediment transport associated with these differences in velocity. These flow induced differences will require substantial "calibration" within the model to yield reasonable estimates of sediment/contaminant flux. Given the non-linearity of both the velocity/shear stress and the shear stress/transport relationships it will be unlikely that such calibration will result in accurate simulations across a wide range of flows. This may be the principal reason that the model, at least some locations, seems to over-predict TSS values at low flows while under-predicting them at high flows. The significance of such variations will tend to increase with the duration of the model run and may become more of a problem during the extended runs planned for the Corrective Measures Study. **(3B)**



Key: Black line — Simulated
Green line — Measured

Figure 6.2-5 Comparison of Measured and Simulated Flow and Elevation at Five Stations During the October 2003 Storm Event

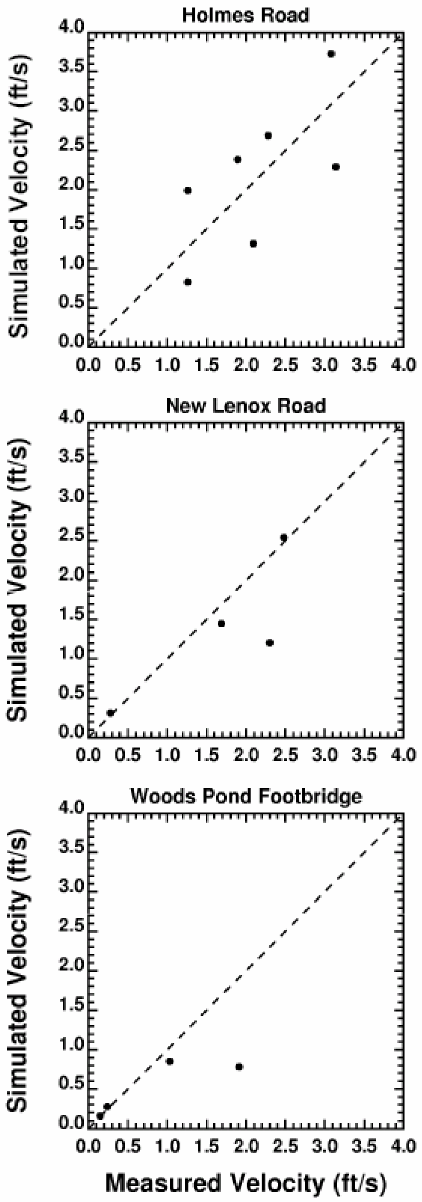


Figure 6.2-8 Comparison of Measured and Simulated Velocities for Holmes Road, New Lenox Road, and Woods Pond Footbridge During the 26-Year Validation Period

▶Comparisons between measured and modeled velocities are complicated by the model use of a single grid cell across the main stem channel. If the single cell is to be retained measurements should be designed to yield cross-sectional averages as opposed to point measurements. The majority of available data do not appear to be suitable for this purpose.

While on the subject of model grid specification, I must repeat what has been stated before regarding the advisability of increasing the number of grid cells across the channel. My review of available data detailing bathymetry and sediment type indicates that accurate simulation of these characteristics and their proper incorporation within transport models requires a minimum of three grid cells rather than the present one. If this proves (or is known) to result in an unacceptable increase in computation time then consideration should be given a reduction in the lateral extent of the floodplain since there are a variety of data (see Example Model Simulations, e.g.) that suggest that there are substantial areas along the inshore limits of the floodplain that are only occasionally affected by flooding and where, as a result, minimal longterm changes are to be expected. Alternatively, a coarser grid might be used on the floodplain. **(3A)**

▶Beyond this matter of flow, velocity and shear stress, I was pleased to see that the current model includes direct calculation of side bank erosion. I'll leave the adequacy of this formulation to those more qualified than I but must state my concern over the apparently simple partitioning of the mass of sediment supplied by this process between surficial erosion (during the rising hydrograph) and mass failure (during falling stage). It's hard for me to see why the masses should in anyway be equivalent. This subject was also noted in GE's review of the Model Validation Report (MVR). Justification for this approach should be carefully presented using the field data and/or supplementary data from previous publications. **(1D)**

▶The side bank issue affects both the margins of the river channel and the floodplain. All indications suggest that this latter area is primarily a sink for sediment and PCBs. As I understand it, the model treats each of the grids on the floodplain in a manner similar to those in the river channel and seeks to erode the soil surface by flow induced shear during flooded conditions. Given the presence of vegetation this very seldom occurs leading to continuing deposition over most of the area. What sediment and PCB that is supplied by the floodplain comes from aperiodic failure of the floodplain margin or sidebank.

If correct, this scheme seemingly neglects any transport associated with the movement of leaf litter and/or the rainfall induced wash-off of materials adhering to the surfaces of vegetation following flood inundation. Has consideration been given to the inclusion of these factors in the model? If not, is there a solid basis for their neglect? It may be that this is a subject that could be quantified in the revised monitoring program recommended above. It may also be that phytoremediation should be included in this program (if it has not previously been investigated) and/or included in the upcoming CMS. **(3C)**

▶Moving to the areas of standing water in the main channel, the backwaters and Woods Pond, I remain concerned about the accuracy of the sediment transport formulations. This concern is not entirely alleviated by the sensitivity analyses presented in the previous calibration report (Weston, 2004) and those included in the MVR. The response to a 50% variation in a variety of parameters overall seems reasonable and makes clear that all of the model results are sensitive to upstream boundary conditions. This, of course, is not surprising given the role of streamflow across this boundary in model dynamics and the limited number of areas within the PSA sufficient to serve as a sediment supply. What's demonstrated is in fact why one worries about the accuracy of the sediment transport formulation. **(1)**

▶Nor are concerns alleviated by model runs requiring what appear to be inordinately high diffusive fluxes to explain the simulated increase in PCBs at New Lenox road relative to those at Holmes Road (see pp.6-72 and Fig.6.2-42 MVR) during low flow conditions. This response suggests that the calibration of the sediment transport formulation might, because the majority of the available data were obtained during high flow events, have produced a function that is overly specific to higher flow conditions. This calls to mind the comments of Dr. Lick regarding the need to reduce the number of adjustable parameters in the transport formulation and more accurately specify those that remain (comments that I second) so as to have confidence that the algorithm is an accurate representation of governing physics and not simply some curve fitting routine.

I would recommend that increased focus be placed on the sediment transport formulation and that model runs be conducted in which the sediment supply across the upstream boundary is set to zero. Minimal " tuning" and reasonable results under these conditions would increase confidence in the formulation and directly benefit future evaluations in the CMS that very likely will be dealing with transport in specific areas within the PSA and require accurate estimates of local mass movements. **(1)**

▶Does the presence and movement of ice seasonally contribute to sediment erosion in the PSA? Within the channels or along the flood plain? **(13)**

2. Are the comparisons of the model predictions with data sufficient to evaluate the capability of the model on the spatial and temporal scales of the final calibration and validation?

▶This depends to some extent on the characteristic being studied. For stream flow and stage model/data comparisons are based on a relatively long data set covering a wide range of seasonal, annual and intra-annual conditions at a number of sites throughout the PSA. The resulting comparisons are clearly sufficient to evaluate model capabilities over a relatively long period of time.

Moving to velocities and ultimately shear stress involves a significantly shorter data set at a limited number of locations. This does not necessarily mean that the data are inadequate since these characteristics are not expected to significantly change with time. As discussed above, however, it is the comparisons presented in the MVR that are less than sufficient. Better use of the available data would be the place to start. Careful review of the results of these analyses may then point to the need for additional data from differing locations and/or modified measurement procedures. **(3B)**

▶The TSS data set appears to be moderately robust although I worry that too much emphasis might have been placed on storm conditions. The data also appear to be primarily point measurements with some representing integrated measurements over the vertical. The absence of time series data prevents detailing of processes such as the time scales of resuspension and deposition during rising or falling stage with particular emphasis on the onset of resuspension during relatively low flow conditions. The influence of such low level but persistent resuspension and transport on PCB fluxes is largely ignored in the present study which places primary emphasis on storm events. Absent the low flow details its difficult to assess the role of these processes in the long term. Such assessments will be a subject of study in the upcoming CMS. As noted above I'd recommend immediate initiation of a monitoring program designed to provide time series observations of TSS concentrations at a number of selected stations throughout the PSA. **(1A, 13)**

▶With regard to PCBs the data set allows at least an initial evaluation of the spatial and temporal validity of the model. Specification of the initial concentrations used in the Validation remains a concern since to some extent the hind cast method used to develop initial concentrations is not entirely independent of the subsequent model run. This apparently was required by the limited data available for 1979. Review of the RCRA Facility Investigation Report (BBL, 2003) shows the results of sampling dating back to the 1979-1980 period. It seems possible to use this variety of data to provide at least a check on the trends and associated initial concentrations suggested by use of the model. Might this be possible? I'm assuming that extensive "data mining" has been part of this exercise and that therefore some of this approach might already have been tried. If so, a brief discussion of this in the report would be useful. **(7)**

▶The model/data comparisons for PCBs should be extended to include consideration of distributions over the vertical. Any assumption of a well mixed layer extending over depths in excess of a few inches doesn't seem to agree with field data (see e.g. Fig. 4-21c, BBL,2003, attached). How are these differences to be reconciled? (i.e. use of 6" well mixed layer vs. detailed core data sowing little mixing beyond 1-2in). These core data and associated radio-dating also allow estimates of sedimentation rate to be compared to the mass flux data provided by the model. This comparison was not part of the validation report. It would provide an additional check on model results and is recommended. **(3D)**

▶An additional check on model results that should be considered includes the use of time series bathymetric data to check on the accuracy of sediment erosion/deposition estimates.**(1A)**
▶Examination of these data might also be part the studies dealing with the recommended increase in the number of model grid cells across the river channel. **(3A)**

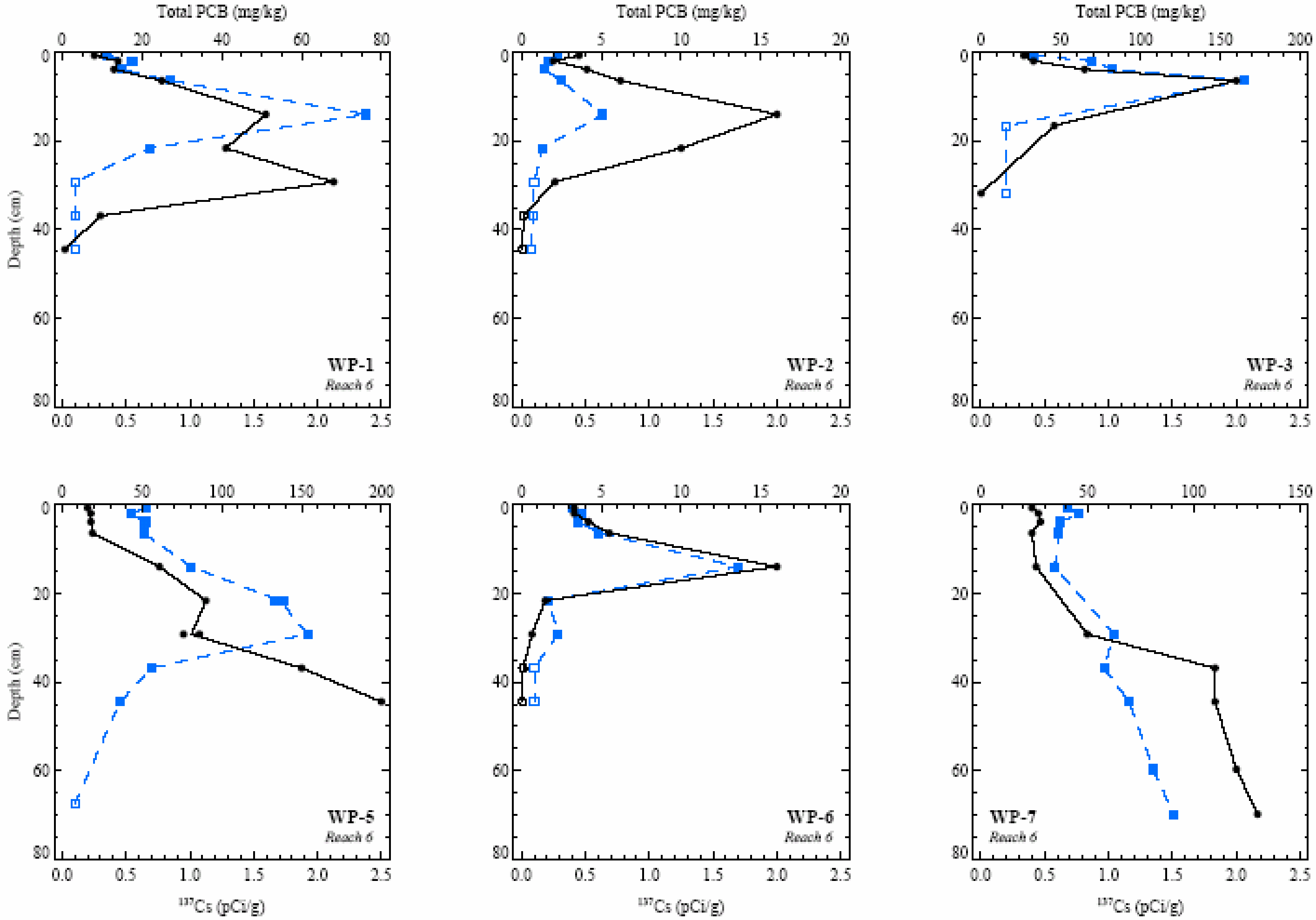


Figure 4-21c. Depth profiles of GE high resolution cores collected within Woods Pond during 1994 - 1995.

Note: Cesium-137 plots are all on the same scale; PCB scales vary by core. Line represents average of duplicate samples.

IF - I:\GENhou\documents\reports\RFI_report\Section4\figures\Fig4_21c__GE_high_res_cores_WP.pro
Thu Jun 19 16:06:08 2003

▶The above comments refer only to the PSA. It would appear that there are insufficient data to adequately test model results in both the upstream (which is in a state of major change) and the downstream model. Efforts to gather these data should be initiated immediately. **(8, 9)**

3. Is there evidence of bias in the models, as indicated by the distribution of residuals of model/data comparisons?

▶As noted on several occasions above, visual examination of the figures suggests that the model overpredicts TSS concentrations at low flows and underpredicts them at high flows. The distribution of residuals provides clear indication of the overprediction at low flows for most stations (see Table 6.2-6). Although a number of reasons for these differences (both high and low flow) are presented there is no mention of the possibility that they are the result of less than accurate formulation of the sediment transport process, in particular bed erosion/deposition, in the model. As discussed above, this seems more than possible and should be carefully evaluated since it will be of increasing concern during the CMS. **(1A)**

▶In terms of PCBs, the model appears to overpredict concentrations in Woods Pond (see Fig.6.2-50). This is likely the result of inaccurate specification of suspended sediment flux associated with the above TSS predictions. Alternatively it may be associated with the specification of boundary conditions and/or the model treatment of PCB sequestering in Woods Pond. The model decline within the Pond appears to be approximately linear with time while the data suggest an exponential distribution. The latter might be the result of progressive source control (both natural due to burial and anthpogenic) acting in combination with sedimentation of cleaner materials. The model might easily obscure this process since it assumes a well mixed sediment column to 6in. With sedimentation rates in Woods Pond averaging approximately 2-3 mm/yr the above depositional processes might be expected to influence the upper 3in over 25 years. This might very well result in a reduction in actual PCB concentrations that would not be well simulated by the model. **(2A)**

4. Have the sensitivities of the models to the parameterization of the significant state and process variables been adequately characterized?

▶In most cases the sensitivity of each of the models to the significant state and process variables has been adequately characterized. However, as discussed above, I remain concerned about the accuracy of the sediment transport processes within the model domain and look to the sensitivity analysis for some guidance. The sensitivity analyses provided in the MVR and the earlier calibration report provide clear indication that sediment mass flux throughout the PSA is very sensitive to the upstream boundary conditions. With these reasonably well specified, using a combination of HSPF results and field observations, it may not be very difficult to achieve reasonable model/data agreement despite inaccurate simulation of bed erosion within the PSA. The sensitivity analyses do not take the next step to test the adequacy of the interior formulations by shutting off the boundary contribution of sediment and PCB’s. The importance of this formulation will progressively increase as the upstream sources of contamination are reduced and remedial measures for the downstream areas are being evaluated. A detailed analysis of the sensitivity and accuracy of this formulation is recommended The results of this effort might serve to address the concerns raised in GE's review of the MVR dealing with the relative importance of Reach 5A and 5B as sources of sediments and contaminants to Reach 5. **(5A)**

▶In addition to testing of the erosion formulation additional sensitivity analyses should be performed to assess model response to the thickness of the active sediment column. This has been a matter of concern for some time. Experience in other riverine/bayou systems as well as the detailed core data (Fig. 4-21c, attached) indicate to me that the active bed used in the model is too thick. It may be that this specification in terms of physical transport characteristics has relatively little effect on overall model results (although that might bring up another set of questions). A test of model response to this characterization is recommended. **(3D, 5A)**

5. Are the uncertainties in model output(s) acknowledged and described?

▶The uncertainty analyses presented in the MVR are interesting and represent a real attempt to develop new methodology for the assessment of complex models such as EFDC. Both Kolmogorov-Smirnov (KS) and Response Surface Modeling (RSM) analyses were applied. The KS methods require numerous model runs to develop a statistically robust data set. Given the time required for each run this was an onerous task and one made no easier by the failure of four of the EFDC runs. This effort yielded an initial estimate of uncertainties in the model prediction of PCB concentrations both water column and sediment associated. These concentrations were emphasized since they are passed to the food chain model (FCM). The KS analysis was supplemented by RSM to evaluate model uncertainty due to uncertainty in input parameters including flows, roughness parameters, critical erosion stresses, etc. The evaluations include consideration of the effects of the calculated uncertainties on the FCM predictions.

Overall the results of the uncertainty analyses, while interesting, were difficult to interpret in terms of model abilities to accurately simulate PCB transport and fate. The presentation of a series of individual summary figures for each condition does not facilitate interpretation in the absence of detailed description of cause and effect. The majority of the supporting text dealt primarily with methodolgy rather than interpretation. It may be that this type of analysis and its sophistication is premature and requires a greater understanding of the processes included in the model and their interactions than we presently have. As this becomes available with model use a better description and analysis of uncertainty might be possible.

Moving from the formal analyses of uncertainty to the general subject of model uncertainty, this report (MVR) too often fails to provide detailed discussion of uncertainty including consideration of causes. Statements such as "The simulated hydrographs....reproduced measured hydrographs reasonably well, however in some cases the magnitude of the simulated flow differed from the data in both positive and negative directions" (see pp 6-35) are too common. Uncertainty is to be expected as is variability both due to the input data being used and numerical model response. It should be introduced early in model discussions and included in logical discussion of cause and effect relationships throughout the report. I would hope that any summary report considers this as an absolute necessity. It will be of great value in building the confidence of a general readership. **(5B)**

6. Upon review of the model projections of changes in PCB concentrations in environmental media in the example scenarios, are such projections reasonable, using your technical judgement, and are they plausible given the patterns observed in the data?

▶The model response for the two examples shown seems reasonable in the sense that most locations experience some decline in PCB concentrations over the upper 6in of the sediment column. I would have expected to see more of a change in Example 2 since virtually all of the major sources of PCB to the system have been shut off. This may point to the importance of the floodplain as a continuing source of PCBs to the system and/or point to the fact that chage can only be expected to occur over depositional depths. These depths, controlled by sediment rates in the area, will seldom exceed 3in except in the vicinity of point bars or similar channel features that are not well modeled because of the coarse grid scales used in the river channel. Since the model is averaging over a significantly greater depth (6in) the changes over 26 years may be difficult to assess. I'd be interested in seeing what a model run of 50 years or more would show. **(10)**

7.Is the final model framework, as calibrated and validated, adequate to achieve the goal of the modeling study to simulate future conditions 1) in the absence of remediation and 2) for use in evaluating the effectiveness of remedial alternatives ?

▶This model system is not yet ready for use in the CMS. Several issues remain to be addressed. First, as discussed above, the hydrodynamics must be more thoroughly verified so as to insure accurate specification of boundary shear stresses. This factor will be central to future evaluations of selected remedial schemes such as capping and is presently essential within evaluations of sediment and PCB transport. The information and data provided in the validation report does little to build confidence in the present formulation. **(3B)**

▶Next, additional work is required to develop an accurate formulation of sediment transport. The suggestions of Dr. Lick in terms of both equation parameters and the structure of the sediment bed should be carefully evaluated. There seems to be an abundance of data and experience to suggest that 6in is an overestimate of the active bed thickness. There is also concern that the model as it exists may be biased to high flow conditions. Add to these questions regarding side bank erosion and floodplain dynamics. **(1, 3C, 3D)**

▶Even with these process questions resolved there remains the issue of run-time. It seems clear that the model as presently configured requires entirely too much time for the completion of a single run to be useful within timely evaluations of a significant number of remedial schemes. We probably knew this several years ago and should have been more sensitive to the need to develop alternative formulations. A number of these, including the separation of the hydrodynamics from the transport estimates and subsequent FCM evaluations were previously mentioned. It is now necessary and possible to go further. Using the experience gained from “whole model” runs it should now be possible to develop a number of synthetic hydrographs detailing streamflow, stage, and TSS concentrations at the upstream boundary and each of the primary tributary streams. This would eliminate the need to run HSPF on a regular basis. EFDC is a ponderous model and can be streamlined now that we have a better idea of the relative importance of the governing variables. John Hamrick should be charged with this task (no more than 6 months) as soon as possible. As part of this streamlining the model grid characteristics should be carefully reviewed, again using what has been learned about the relative importance of each of the domain regions (sidebanks, backwaters, floodplains etc.). My sense of the present grid is that it underspecifies the channel region and overspecifies the floodplain. The latter could almost be treated as a box with fluxes simply proportional to stage which could be specified along its margin by EFDC. If in time more detail of the interior of the plain is needed for remedial purposes consideration might be given to replacing the high resolution grid on the plain while placing a box in some other area (channel sidebanks?). **(3A, 4, 13)** Finally, I’d consider eliminating the FCM in favor of a parametric (flow, TSS concentrations?) approximation of body burden uptake based on the results of the complete model runs. **(3A, 4, 11, 13)**

▶In short, what I’m thinking about is the development of an supplementary modeling scheme for use in the CMS. The complete model would serve as a guide assisting in the development of a series of simpler, more efficient but less comprehensive, formulations that would be directed at particular remedial schemes. The complete model framework would remain in place providing guidance regarding the need for and type of data to supplement model formulations for both calibration and verification purposes but would not be run as frequently as the supplementary schema. The alternative might be to turn to a different series of models entirely. This is not recommended without good reasons that I don’t have at the moment. **(4, 13)**

Personal Comment

With all these reviews it’s easy to loose sight of the amount of work and dedication that it has taken EPA and GE and their contractors to get us to this point. As indicated earlier this has been and continues to be an ambitious project dealing with a complex subject. You have made significant progress and in many cases developed new methodology that will benefit future investigators. You are to be complimented on your dedication, skill, and patience. Stay the course! This can be accomplished.

References

BBL 2003 Housatonic River-Rest of River RCRA Facility Investigation Report. prepared for General Electric Company, Pittsfield, Ma. by Blasland,Bouck & Lee Inc and Qualitative Environmental analysis, LLC (QEA).

EPA 2006 EPA Response to Questions From the Model Validation Document Overview Meeting. DCN:GE-061406-ADFI June, 2006

Weston 2004 Model Calibration: Modeling Study of PCB Contamination in the Housatonic River. Appendix B Hydrodynamic and Sediment/ PCB Fate and Transport Model Calibration. Prepared for U.S. Army Corps of Engineers, Boston,Ma. and U.S. Environmental Protection Agency by Weston Solutions Inc. West Chester, Pa. DCN:GE122304-ACMG.

Weston 2006 Model Calibration:Modeling Study of PCB Contamination in the Housatonic River. Vol 1 and Vol 2. Prepared for U.S. Army Corps of Engineers, Boston,Ma. and U.S. Environmental Protection Agency by Weston Solutions Inc. West Chester, Pa. DCN:GE-030706-ADBR

APPENDIX A.3

COMMENTS OF DOUGLAS ENDICOTT

PEER REVIEW COMMENTS ON
THE MODEL VALIDATION REPORT
Douglas Endicott

Introduction

I appreciate the opportunity to review the work of the modeling team for the Housatonic River. This has been an interesting learning experience for me. I would also like to compliment the modeling team for having conducted this project in a thoroughly professional manner. In addition, I appreciate their responsiveness in the short time frames allowed for the peer review process.

▶I do have several comments to make regarding the peer review process itself. In the *EPA Response to Questions from Model Validation Document Overview Meeting*, EPA correctly notes that models are uncertain and that a degree of uncertainty is inherent in “these types of studies”(question: What scientific study does not include a degree of uncertainty?). They go on to identify three principal sources of model uncertainty (models are simplifications/approximations of reality, parameters are uncertain, future conditions are unknowable). Although these can be significant sources of uncertainty, EPA chose not to highlight several others, which may be equally important:

- Inappropriate process representations/formulations,
- Errors resulting from discretization of the model or the scale of its application,
- Errors in programming, preprocessing, running, linking and postprocessing the models, and
- Untested model predictions.

These are “lurking” sources of model uncertainty, which are not typically revealed by uncertainty analysis. I think it is useful to acknowledge these pitfalls, and keep them in mind throughout the course of a project. I think much of the Peer Review deliberations, and more than a few of the comments, reflect concern about these factors. **(5B)**

▶It is unfortunate that the peer review process was conducted in more of a confrontational, as opposed to a collaborative manner. The scientific process is not well served by the former approach. I think more would have been accomplished in the peer review sessions, had a freer dialog and more open exchange of information and ideas between the various parties been possible. Because of the constraints of the Consent Decree, this was not allowed. As a result, the relationship between the modeling team and the peer review panel was sometimes adversarial, and our recommendations rebutted inappropriately (from my view). Although this may be as the lawyers intended, it also resulted in some lingering issues that have not been appropriately resolved through the entire peer review process. **(13)**

Questions for the Model Validation Report

1. Considering the changes implemented in the Phase 2 Calibration, does the model as calibrated and validated, based on your technical judgment, reasonably account for the relevant processes affecting PCB fate, transport, and bioaccumulation in the Housatonic River to a degree consistent with achieving the goal of the modeling study?

The modeling frameworks (HSPF, EFDC and FCN) applied in the Housatonic River are state-of-the-art. Within each model, all relevant processes are represented in a conventional manner. Therefore, I conclude that the models reasonably account for the relevant PCB transport, and bioaccumulation processes.

Of course, this question is somewhat ridiculous because a successful model must do more than just "account" for processes! Instead, it is necessary to reevaluate the modeling framework, its implementation, and its parameterization in light of our understanding of PCB fate, transport, and bioaccumulation processes as well as the behavior of the Housatonic River as an ecosystem. Two primary criticisms emerge from this reevaluation:

1. ▶EFDC fails to address lateral variation in hydrodynamics and sediment transport processes, and **(3A)**
2. ▶There appear to be parameterization errors in specific sediment transport and PCB fate and transport processes. **(1)**

▶I should note here that, since the upstream and downstream models share the same frameworks, the same comments apply. **(8, 9)**

▶On the first point, EFDC appears to fail to address lateral variation in hydrodynamics and sediment transport processes, because the model grid in the river channel is one-dimensional over most of the Rest-of-River (ROR). Prior applications of SEDZL (the model upon which the EFDC sediment transport code was apparently based) to river systems have used a 2-dimensional, vertically-integrated grid, using 5 to 10 lateral grid cells to represent the channel cross section. The 2-dimensional sediment transport models have been demonstrated to predict scour and deposition in river channels with reasonable accuracy[1]. These applications have demonstrated that lateral variation in current velocities and bottom shear stresses result in considerable variation in sediment scour depending upon location in the channel. This behavior is often observed in natural channels as a pattern of erosion from mid-channel/deposition nearshore, or erosion on the outside of a bend/deposition on the inside. Because the relationships between current velocity, shear stress and sediment erosion are nonlinear, it appears likely that sediment scour

---

[1] See: Gailani, J.Z., W. Lick, K. Ziegler, and D. Endicott. 1996. Development and Calibration of a Fine-Grained Sediment Transport Model for the Buffalo River. *J. Great Lakes Res.,* 22(3):765-778.

predicted by the 1-dimensional model of sediment transport in the Housatonic River will be different than the scour predicted by a 2-dimensional model. How *much* different is unknown, because the EPA modeling team have not tested the sensitivity of the EFDC sediment transport model to the number of lateral grid cells in the river channel. This is a significant issue, because release of PCBs from eroding sediments is a major source of the contaminant in the ROR. **(3A)**

On the second point, there appear to be parameterization errors in specific sediment transport and PCB fate and transport processes. These include the parameterization of sediment scour and deposition, the thickness of vertical sediment segments and rate of bulk mixing between these segments, and the diffusive exchange between sediment pore water and the overlying water column. In addition, the appropriateness of the equilibrium partitioning assumption for PCBs on resuspended sediments has been repeatedly questioned. These issues are discussed below. Where there appears to be a tie-in between parameterization errors and lack-of-fit or bias in model predictions, this will be noted.

- ▶The vertical representation of the sediment bed in the PCB fate and transport model: The thickness of the surficial sediment layers have been specified as either 4 cm (Reach 5A) or 7 cm (Reaches 5B thru 6). These layers are assumed to be well-mixed, via a combination of physical and biological processes. In addition, sediments from the surficial layers are mixed with deeper layers, down to a depth of 30 cm below the sediment-water interface. Analysis of benthic invertebrate data is offered to justify these assumptions. However, these data show that below Reach 5A the benthos are predominantly filter feeders or surficial sediment feeders/dwellers. Such organisms are unlikely to cause sediment mixing below the top couple of centimeters. 7 cm of surface sediment mixing is too deep, and mixing of sediment between the surface layer and deeper layers is inconsistent with the life history data for the resident benthos in these reaches. The consequence of too much sediment mixing (and too-thick surficial layers) is that the PCB fate and transport model will predict an unreasonably slow rate of decline in surficial sediment PCB concentrations, which is observed in both Phase 2 calibration and validation periods. **(3D)**
- ▶Sediment transport parameterization: Cohesive sediment resuspension is parameterized based on the analysis of data from SEDFLUME experiments conducted on a number of sediment cores. Results of this analysis produced shear stress exponents of 1.59 for sediments in Reaches 5A and 5B and 0.95 in Reaches 5C and 6[2]. These values are considerably lower than shear stress exponents reported in the literature. For example, Lick, Ziegler and coworkers have reported that for cohesive sediments tested at a significant number of sites, the shear stress exponent is generally constrained within a fairly narrow range (n=2.6±0.3). Based upon this, it appears that parameterization of

---

[2] Model Calibration Report, Volume 3 - Appendix B Hydrodynamic and Sediment/PCB Fate and Transport Model Calibration, Attachment B.5 Analysis of Sediment Erosion Data

cohesive sediment resuspension in the model is erroneous, and should be corrected consistent with the guidance offered by Dr. Lick. As a consequence, it will also be necessary to recalibrate deposition rates. **(1A)**

- ▶The diffusive flux (i.e. sediment-water exchange) of PCBs from sediment pore water to the overlying water column was parameterized as a constant mass transfer coefficient during Phase 1 calibration. Phase 2 calibration has altered the spatial profiles of PCB water column concentrations at low flow conditions, suggesting that sediment-water exchange should be recalibrated. In all likelihood, the mass transfer coefficient should be increased, and made a reach-specific parameter via a relationship between the mass transfer coefficient and the benthos density. **(1B)**
- ▶EFDC uses the same equilibrium partitioning (EP) model as every other fate and transport model in use today. EP generally works quite well, and simplifies the model computation by requiring only a single state variable for each chemical. It has been suggested that desorption kinetics should be incorporated into the Housatonic River model, because PCBs desorbing from resuspended sediment might not reach equilibrium within the timeframe they remain in suspension. This disequilibrium may help explain the pattern of PCBs moving in “ribbons” of sediment downstream, and the high surficial PCB concentrations in Woods Pond sediment. I think this is mostly speculation and, although I am curious to see if it is true, it is well and beyond the objectives and goals of this modeling study. **(1C, 2B)**
- ▶I have already commented on the misattribution of the assumption made regarding PCB partitioning to non-filterable organic carbon in the water column ($K_{doc} = K_{poc}/100$), and EPA has responded (less than satisfactorily) to this issue. Regarding the parameterization of the 3-phase equilibrium partitioning model for PCBs, a useful citation (i.e., the Burkhard references) is: USEPA, 2003. *Methodology for Deriving Ambient Water Quality Criteria for the Protection of Human Health (2000), Technical Support Document Volume 2: Development of National Bioaccumulation Factors.* United States Environmental Protection Agency, Office of Water, Office of Science and Technology. EPA-822-R-03-03 (www.epa.gov/waterscience/criteria/humanhealth/method/tsdvol2.pdf) **(1C)**

▶This question also asks us to consider changes to the models resulting from additional (Phase 2) calibration. This specifically applies to EFDC calibration, since little or no changes were made to either the HSPF or FCN models. Previously, the Peer Review panel commented that the 14-month calibration was too short to reflect many significant long-term model processes, specifically the evolution of PCB concentrations in the sediment bed. So the EPA modeling team went back and extended the calibration to cover a much longer, 10-year duration. I expected to see some changes in the model parameters related to the residence time of PCBs in the surficial sediment layer. Instead, as a result of Phase 2 calibration we see some dramatic changes in the simulation of PCB transport and fate in the Housatonic River: There are now 3 times more PCBs entering the Primary Study Area (PSA) at the upstream boundary, although the difference (compared to the Phase 1 calibration & flux analysis) diminishes further downstream (2 times more PCBs at New Lenox Road, 1.5 times more at Woods Pond). Export of in-place PCBs from the sediments of Reach 5A is no longer the predominant contaminant source to the Rest-of-River. To me, this is a breathtaking change in the PCB fate and transport simulation, which rewrites the entire assessment going back all the way to the conceptual model presented in the MFD. I was disturbed to see such a major change taking place so late in the model development process, and was also somewhat surprised at the nonchalance with which this revision was presented and received.

Does the hydrograph or the longer duration of the Phase 2 calibration explain these differences? To some degree, maybe. In 10-year period, there are a number of high-flow years (e.g., '90, '96,'98 and '00) as well as low-flow years ('91, '92 and '99). The hydrograph impacts the balance of sediment erosion and deposition averaged over different durations. The MVR does not discuss this to any significant extent. However, the change in the PCB fluxes mostly reflects a new treatment of the boundary condition, apparently provoked by some new boundary condition data. Comments regarding the boundary condition used in the Phase 2 calibration are provided in my response to question 3. At this point, EPA should consider whether further refinements and improvements to the model will occur in the future, as more data are collected and the models are run. In other words, are the models only as good as the latest data point? How can the model calibration be considered "final" if, as the Phase 2 recalibration demonstrates, a few new data can so readily change the PCB flux analysis and possibly other model results? **(6)**

▶Another change that has been undertaken by EPA is to extend the spatial domain of the models to cover the upstream reach of the East Branch through the area that has been remediated, as well as downstream from Woods Pond dam to Rising Pond dam. It is encouraging to see this development, because the modeling domain confined to the ROR was clearly too small to fully address the extent of PCB contamination in the Housatonic River. However, I am not convinced that either the upstream or downstream models are adequately calibrated and validated at this time. It is obvious that much less effort has gone into these models, and the presentations made by the EPA modeling team at the Document Review Meeting looked much more like progress reports than finished products. **(8, 9)**

2. Are the comparisons of the model predictions with data sufficient to evaluate the capability of the model on the spatial and temporal scales of the final calibration and validation?

Graphical and statistical comparisons between model predictions and data are the most objective measures of model performance. In general, sufficient comparisons have been presented in the MVR and supplemental material with which to evaluate the capabilities of each of the models. Specific comments are offered below, regarding the strengths and weaknesses of the various prediction-data comparisons, as well as identifying several comparisons that have not been made, but should be. In addition, errors in model predictions are revealed by some of the comparisons, and these are pointed out so that they can be corrected by the modeling team.

▶EFDC predictions of flow, TSS and water column total PCB concentrations were compared to data using both longitudinal profile and time series plots. Although the spatial profiles for TSS predictions appear reasonable for several selected days (Figures 4.2-34 thru 4.2-37), the time series plots indicate significant and persistent overprediction of low TSS measurements. This is particularly evident at Holmes Road (Figure 4.2-38), but also at New Lenox Road (4.2-39) and to a lesser degree at Woods Pond headwater (4.2-40). In the MVR (page 4-59), EPA offers that "Simulated TSS concentrations pass through approximately one-third of the data…", which is surprisingly poor model performance. Similar overprediction of low TSS measurements is apparent for the validation period as well. According to comments by both EPA and GE, the overprediction of low TSS measurements is due to significant bias in the specification of the East Branch boundary condition on days when TSS concentrations were not measured. Time series plots for specific flow events (Figures 4.2-42 thru 4.2-48) indicate a better fit of TSS predictions under high flow conditions. **(6)**

▶Sedimentation rates predicted by EFDC for individual grid cells in the river channel over the 10-year Phase 2 calibration period are plotted in Figure 4.2-58. This figure shows a highly variable pattern of deposition and erosion, especially in Reach 5A, where deposition fluxes exceeding 5 cm/yr are predicted for cells in close proximity to others in which erosion fluxes of 1-2 cm/yr are predicted. Predicted sedimentation rates are not tested by any data at this scale, however. On a reach basis, relatively small deposition fluxes (0.2-0.5 cm/yr) are predicted throughout the ROR for both the calibration and validation periods. Figure 4.2-60 graphs the sedimentation rate in grid cells within Woods Pond, along with sedimentation rates measured in cores. There are many EFDC grid cells in Woods Pond, and a significant number of dated sediment cores. I specifically examined the 1995 sediment core data suggested to be most representative in terms of cesium-137 profiles, and found that both measured and predicted sedimentation rates are in the range of 1 to 10 mm/y. On a pond-wide average basis, sedimentation rates predicted for the calibration period agree fairly well with the sediment core data (Table 4.2-5). Over the 26-year validation period, the predicted sedimentation rates appear to be somewhat low in comparison to data (Table 6.2-7).

Predicted sedimentation rates were also compared to measurements of bathymetric change made between 2000 and 2006 at a series of transects in Reach 5A (Figure 6.2-33). The measurements of bed elevation change (positive and negative) were consistently larger and more variable than the corresponding model prediction although, since the predictions and data were both spatially averaged, I have trouble interpretting the significance of this comparison. As with other EFDC results for the sediment bed, grid cell predictions were aggregated to mile or sub-mile intervals. At finer spatial scales, there appear to be problems with the fidelity of the model. For example, the transect data from location XS153 show several feet of deposition while EFDC predicts significant erosion at this location, both during calibration (erosion=1”) and validation (4”) periods. The transect data themselves (presented at the Document Overview meeting: *09 - Erosion& Deposition - Garland.pdf*) show relative nonuniformity in sedimentation both longitudinally (between transects) and laterally (within transects). To me, the data show that lateral resolution is warranted for hydrodynamics (shear stresses) and sediment transport. Other subgrid features (lateral variations in bathymetry, bends, trees in river) are also significant physical features, which are not represented in the model. **(2C)**

▶Model predictions of total PCB concentrations in the Housatonic River water column were compared to data using time series plots and longitudinal profiles. The time series plots (Figures 4.2-65 thru 4.2-68 and 6.2-41 thru 6.2-44) indicate that, for PCB concentrations, there are significant gaps in data collection over the Phase 2 calibration and validation periods. As was the case for TSS, the model tends to overpredict low PCB concentrations, and the overprediction of these two state variables appear to coincide. This is most apparent at the Holmes Road station. Again, this is primarily due to significant bias in the specification of the East Branch boundary condition on days when PCB concentrations were not measured. Inspection of the timeseries of East River PCB boundary conditions (Figure 4.2-64) reveals that the only time East River PCB boundary concentrations are lower than 60 ug/L is when data have been measured below this value, yet probably more than half of the data fall below this concentration. There are sufficient PCB data to evaluate the storm event predictions on only one occasion (May 19-21, 1999 at New Lenox Road; Figure 4-2.76); the model prediction timeseries agrees well with this data. The spatial profiles for PCB water column predictions appear to be quite variable. **(2A)** ▶The low-flow simulations have been criticized by GE as underpredicting the increase in PCB concentrations observed between the upstream boundary condition and New Lenox Road, as well as the decrease in PCB concentrations across Woods Pond. In addition, the profiles show a consistent “drop” in PCB concentrations immediately below the upper boundary. These features have been interpreted as suggesting that processes related to PCB fluxes between the water and sediment may be misrepresented or miscalibrated. **(4)** ▶In the *EPA Response to Questions from Model Validation Document Overview Meeting,* a statistical test (overlapping 95% confidence limits) is offered to show that the apparent decrease in PCB concentrations across Woods Pond may reflect sampling variability. It may be useful to use a statistical test to evaluate the gradient in PCB concentration data between two river locations; however, wouldn’t it be better (more powerful) to test whether the *difference* in PCB concentrations between locations is significantly different than zero?

Timeseries plots comparing the predictions of PCB concentrations in the top 6” of sediment in mile reaches to data are shown in Figure 4.2-77 for the Phase 2 calibration period, and Figures 6.2-49 for the validation period. The predictions appear to exceed the majority of data at the end of both the calibration and validation periods, in some or most mile reaches. Although data prior to 1999 are not abundant, the model also appears to underpredict the rate of PCB sediment decline in many mile reaches, including Woods Pond (Figure 6.2-50) as suggested by the overprediction of the majority of 1999 data. The failure of the model to reproduce the rate of decline in sediment PCB concentrations is a major problem, since in the long-term, PCB concentrations in biota will respond to this rate of change. The MVR fails to address this discrepancy and it's possible explanations, even though this error is evident in both calibration and validation periods. It would also be useful to compare sediment PCB concentration predictions to data for deeper, sub-surficial layers. Slower changes in PCB concentrations would be expected in deeper intervals, reflecting slower transport in deeper sediments. **(2A)** ▶Furthermore, I agree that the approach for developing initial conditions for sediment PCBs (circa 1990 for the Phase 2 calibration and 1979 for the validation) precludes a "robust test" of the long-term model predictions. Consequently, these initial conditions should become part of sensitivity and uncertainty analyses. **(7)** ▶Finally, the increase in surficial sediment PCB concentrations over time within the upper-most sediment reach (RM 135.13-134.89) seems problematic and should be corrected. Is this behavior related to sudden drop in water column PCB concentrations downstream of the boundary condition, for example in the way bed load is initiated? Without access to the model, it is impossible to sort this out. **(4)**

▶Figures 4.2-79 and 6.2-51 are intriguing plots of the change in spatial profiles of the grid cell PCB sediment concentrations over the duration of the calibration and validation, respectively, superimposed with the "shotgun pattern" of data for individual sediment PCB samples. Although no interpretation of the model-to-data comparison is offered, it does suggest that the greater than 90% change in sediment PCB concentrations within many sediment grid cell is both plausible and may help explain the spatial variability observed in sediment PCB data throughout the river sediments. In other words, the model simulation suggests that sediment PCB variability arises from contaminated sediments remaining in locations where erosion is not significant, and (relatively) uncontaminated sediments replacing older contaminated sediments in locations where erosion is significant. **(2B)**

▶I remain concerned that we may be falsely confident in the erosion/deposition predictions made by this model, because the model is being applied and validated at relatively coarse spatial scales. Due to practical constraints, the modeling team did not simulate sediment transport at the resolution demonstrated in other successful sediment transport applications (Thompson Island Pool: HydroQual, 1995; Watts Bar Reservoir: Ziegler and Nisbet, 1995; Fox River: Lick et al., 1995; Saginaw River: Lick et al, 1995 and Cardenas and Lick, 1996; Buffalo River: Lick et al., 1995 and Gailani et al., 1996). The unstated assumption is that applying a similar model on a cruder resolution will correctly average the sub-grid scale phenomena, even though some of these phenomena are demonstrably nonlinear. This assumption is not supported by either data comparisons or numerical convergence testing using alternative grid resolutions. Predictions of net sediment accumulation can only be tested for Woods Pond, because that is the only portion of the ROR where a reach-wide average accumulation could be inferred from sediment data. Predictions of net sediment accumulation are untested in other reaches. This is of more than academic interest, because we are asked to believe in the model primarily on the basis of the water column predictions of TSS and PCB concentrations. Yet these only indirectly measure erosion and deposition. Reach 5A, for example, is described as a dynamic environment with rapid exchange of solids and associated PCBs. Within this and other reaches, there are locations where both deposition and erosion are taking place. If aggregated spatially (into either reaches or miles) much of this behavior is “averaged out”. Aren’t there advantages to validating sediment transport predictions without this spatial averaging? In fact, that is the only way it can be done. **(2C)**

▶The food chain model (FCM) predictions should be compared to the mean or other central tendency of the fish species, reach, and age-specific data, because the model is designed to simulate “average” fish according to these categories. This is confirmed by analyses and graphics (e.g., Figure 4.3-8) displaying that the residuals tend to be smaller when the data are averages of 6 or more fish. Unfortunately, many calibration and validation graphics persist in using individual data. Nonetheless, there is overall good agreement between predictions and the averaged total PCB concentration data, with most residuals falling within a factor of 2 of the predicted values. EPA has revised Figure 4.3-7 to use the 95% confidence limits (or ± 2 standard errors) to better quantify the measurement precision, and the FCM predictions fall within these limits. FCM model performance is similar in calibration and validation periods.

Figure 3-9 (Comparison of mean measured biota tissue tPCB concentrations to FCM results [simulation using linked models]) provides a summary illustration of the FCM calibration. PCB concentrations predicted for invertebrates are in good agreement with the D-Net sample data, except for infauna in Reach 5A (which are substantially overpredicted). The model fits for bullhead and sunfish are very good; PCB concentration in these fish vary little between river reaches. Model-to-data comparisons are not as good for suckers, cyprinids and bass, although the PCB concentration data for these species is also considerably more variable. I assume that at least some of the variability of PCB concentrations in fish reflects the wide range of PCB exposure and/or failure of the food chain to spatially average these exposures on a reach basis. If I look for an overall pattern in the residuals on this figure, it would be a tendency for the predicted PCB concentrations to increase moving downstream, which is not reflected by the data. PCBs predicted in Reach 5D for cyprinids, sunfish and bass significantly exceed the measured (as well as predicted) concentrations for these species in other reaches. Since PCB concentrations were not measured in fish from this reach, these predictions are untested. Reach 5D is the only portion of the FCM where predictions based on PCB exposures derived from data meaningfully diverge from predictions based on EFDC exposures. **(11)**

There are insufficient comparisons between data and predictions for both the upstream and downstream models. ▶In the upstream model, there are too few TSS data to tell whether the model predictions are reasonable or not. For both TSS and PCBs, my impression is that the model predictions are much more variable than the few available data. PCB concentrations appear to be substantially overpredicted (e.g., a factor of 5 to 10 overprediction in PCB concentrations). **(8)** ▶In the downstream model, predicted TSS and water column PCB concentrations are in the same ranges as the data at the Rising Pond outlet, but there is not much fidelity beyond that. Initial downstream model PCB predictions included in the MVR clearly indicated a problem with the balance between erosion and deposition in high-gradient reaches. As presented at the Document Review meeting, this has been addressed by adding an additional noncohesive sediment class in the sediment transport simulation **(9)**.

FCM predictions in the downstream model domain are compared to PCB concentrations measured in fish for Reach 7 and Rising Pond. These predictions look favorable overall, although there may be bias in the predictions for specific species. Results appear comparable in data-based and linked exposure predictions. **(9, 11)**

3. Is there evidence of bias in the models, as indicated by the distribution of residuals of model/data comparisons?

▶A number of significant model biases have been identified and discussed above (Question 2). The most significant are:

- The predicted changes in sediment PCB concentrations are slower than observed,
- Spatial gradients in water column PCB concentrations do not match data under low flow conditions, and
- The treatment of the upstream boundary condition results in serious overprediction of low TSS and PCB concentrations. **(2A)**

▶The first two issues appear related to parameterization errors, as noted previously, and should be satisfactorily addressed by the suggested corrections. The bias in TSS and PCB concentrations due to the boundary condition deserves further consideration.

At sampling locations in the upper reaches of the PSA, the model fails to match half to two-thirds of the TSS data. Predicted values at base flows are ~10 mg/L versus data in the 2-5 mg/L range. There is less discrepancy in the lower reaches. In general, the bias is most apparent at lower flow rates. Censoring TSS data <5 mg/L was offered by EPA as a way of correcting the model performance metrics. A similar problem with low-flow bias for PCB predictions is also apparent in the upper reaches of the PSA, and a similar censoring approach was suggested. However, I do not believe it is appropriate to ignore or discount the bias in model predictions under low-flow conditions. About 70% of an "average year" is low flow conditions, and significant PCB uptake by fish coincides with low flow periods. **(6)** It is unclear whether the low-flow PCB bias has led to food chain miscalibration, as suggested by GE, because the food chain model was initially calibrated using data-based exposure concentrations. These would not exhibit the same low-flow bias as exposure concentrations predicted by the model. **(6, 11)** Nevertheless, I am concerned that bias due to the TSS and PCB boundary conditions may result in model errors that have not been addressed. We know very little about how the boundary condition will evolve over time following upstream remediation; hopefully, ongoing monitoring will be used to update and improve upon this situation.

The specification of the upstream PCB boundary condition is an important component of developing a reliable transport and fate model for the Housatonic River, for each of the time intervals over which the model is to be applied (calibration, validation and forecasting). EPA has developed different interpretations of the PCB boundary condition in the Phase 1 and Phase 2 Calibrations, and these differences produce dramatically different outcomes in terms of the PCB flux analysis generated by the model predictions. The PCB boundary condition developed for the Phase 2 calibration applies the partitioning model to estimate particulate and dissolved PCB concentrations from total PCB measurements, which are then independently regressed to flow rate. The PCB phase concentrations are “capped” to prevent the boundary conditions from grossly exceeding the data at high flows. This is a creative approach; I tried something similar in the Fox River. Unfortunately, it’s not clear whether this approach does a reasonable job of describing the boundary condition data. Although this approach may improve the PCB boundary condition at high flows, in comparison to the Phase 1 calibration, at base flow rates the Phase 2 PCB boundary conditions are substantially worse (e.g., Phase 1: ~40 ng/L at base flow vs. Phase 2 : ~100 ng/L). I am not convinced EPA has made the best use of the boundary condition data, and offer some comments below regarding the approach they used to describe this data. Regardless of these issues, it appears that much of the variability seen in PCB concentrations at the upstream boundary cannot be explained, in the approaches taken in either the Phase 1 or Phase 2 calibration, or other alternatives that have probably already been explored. If the TSS and PCB boundary conditions cannot be improved, the best recommendation may be to add a random error component to the boundary conditions in the sensitivity and uncertainty analyses.

The boundary condition data eres provided by EPA in the spreadsheet “East_Branch_Boundary_Data.xls”. Their interpretation of this data, and the descriptive approach used to model the boundary condition, is described in Section 4.1.5 of the Model Validation Report, with further discussion offered in supplemental responses. Figure 4.1-2 (reproduced below) shows the regressions of dissolved and particulate PCB concentrations versus flow that were used to specify the total PCB boundary condition for the EFDC transport and fate model (except for days when measurements were available). I have a few comments to make based on this figure and review of the boundary condition spreadsheet:

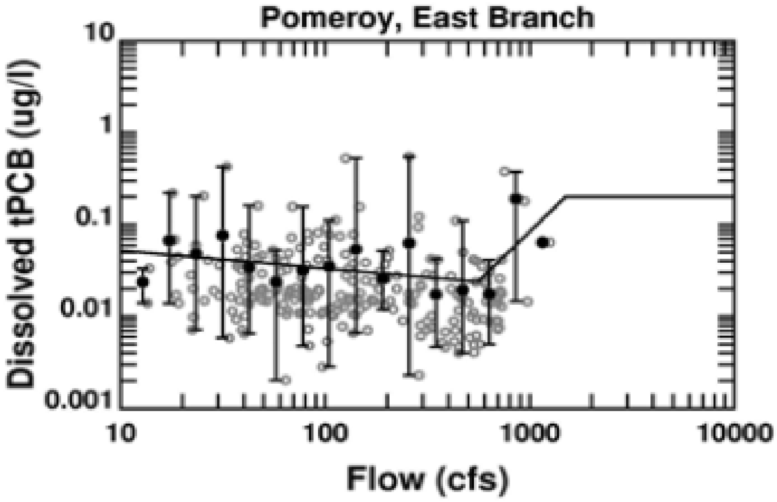
1. As a general observation, I would say that if one focuses on the actual data plotted in these graphs (the open circles), it appears that EPA’s regressions describe relatively little of the variability observed in PCB concentrations at the boundary condition. The regressions are especially tenuous at low and high flows. In fact, the regressions exceed <u>all</u> of the data in the lowest flow “bin”.
2. I cannot find the PCB concentration data collected at flow rates smaller than 24 cfs[3]. Therefore, I have no way to confirm the results shown in Figure 4.1-2 for the two lowest-flow “bins”. At least for the lowest “bin”, the regression line appears to be a poor fit. A noted above, it exceeds all of the measured data.
3. At high flows, the PCB concentrations are “capped” to prevent unrealistically high PCB concentrations during floods. The “cap” was specified as 25 mg/Kg for particulate PCBs, based

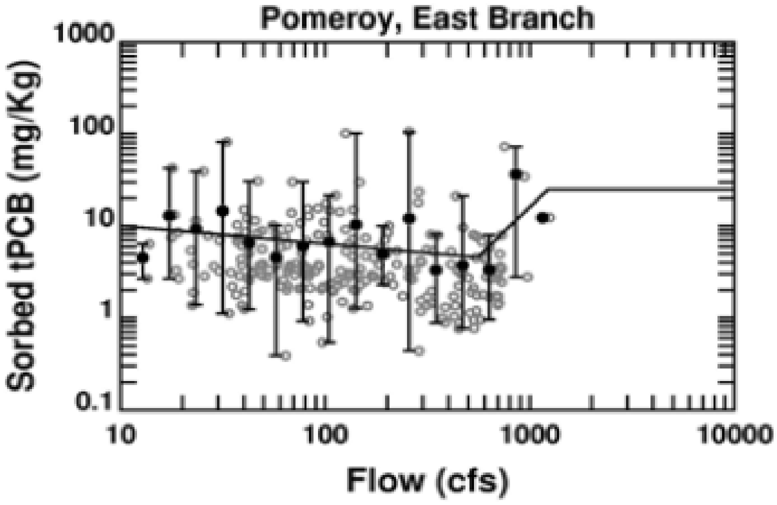
---

[3] I think this may be because East Branch daily flows were provided in the spreadsheet, while Coltsville (hourly?) gauge data were used in the regressions and possibly the graph. I couldn’t find the source of the flow data in the MVR.

upon concentrations measured in sediments deposited on vegetation and banks. I think this value may be somewhat too high. My own calculation of the unbiased mean PCB concentration from the 13 sediment deposit samples is 20.5 mg/Kg, although I was measuring the data off of a graph. Please check this calculation. If 25 mg/Kg is indeed too high, it will bias the high-flow PCB boundary condition.

4. According to EPA, the detection limit was used as replacement values for censored PCB concentrations. While this has little impact on the distribution of total PCB concentrations, it does bias the particulate PCB concentrations that are calculated from total PCB measurements using the partitioning model. I didn't check, but I suspect that dissolved PCB concentrations may be biased as well. I think it is generally preferable to use one-half of the detection limit as replacement values for censored PCB concentrations.





5. Another thing I noticed when I reproduced EPA's partitioning calculation (which is used to estimate the dissolved and particulate PCB concentrations from the total PCB measurements), is that the dissolved PCB concentration estimates don't agree very well with the relatively few dissolved PCB concentrations measured at the boundary condition. This is shown below in Figure 2. If the partitioning model is being used to describe the boundary condition, shouldn't it fit the data?

6. Frankly, some of this data is mystifying. For example, 2 samples were collected on November 12, 1995. For both samples, the flow is reported to be over 1,100 cfs. In the first sample, [TSS] was 450 mg/L and total [PCB] was 0.2 ug/L. In the second sample, [TSS] was 17 mg/L and total [PCB] was 0.52 ug/L. Running these values through the partitioning model, one obtains particulate PCB concentrations of 0.44 and 23 mg/Kg. It is somewhat hard to believe that PCB concentrations on suspended solids could vary by a factor of 50 within one day.

7. It is also curious to me that there has been no shift or evolution in PCB boundary concentrations over the past 16 years, and that seasonality is not evident in the data. If PCB boundary condition data for all years are plotted together, we see the familiar shotgun target (Figure 3). There appears to be some underlying behavior, however, within individual years of data. Figure 4 shows total PCB data for years (especially 2002) in which a positive relationship with flow is evident. Figure 5 shows total PCB data for years (2003 in particular) in which the flow relationship is generally negative. The highest total PCB concentrations in Figure 5 come from the warm (and presumably higher organic carbon[4]) months of July-September, while the lowest PCB concentrations at very high flow were sampled in October and November. While I am confident EPA has looked at this data in great detail, I think the patterns in the data hinted at here suggest some underlying behavior which could be exploited in the regression models used to describe the boundary condition. I believe this might, in turn, provide better PCB boundary condition estimates than the current approach in which all measurements are lumped into one sample.

8. The MVR doesn't discuss what other alternatives may have been tested, but here are some suggestions:
   - AutoBeale (Beal's Stratified Ratio Estimator) is an alternative statistical approach that works well for boundary conditions when the data are stratified by season as well as flow;
   - When regressing the PCB data, were autocorrelation and Lag-1 or 2 considered as factors?
   - Check bias correction formula;
   - Have untransformed PCB data been regressed? How do results compare to log-transformed regressions?

---

[4] EPA is assuming a constant 7.3% organic carbon content on suspended solids, which cannot be confirmed from the data collected at the upstream boundary.

Figure 2

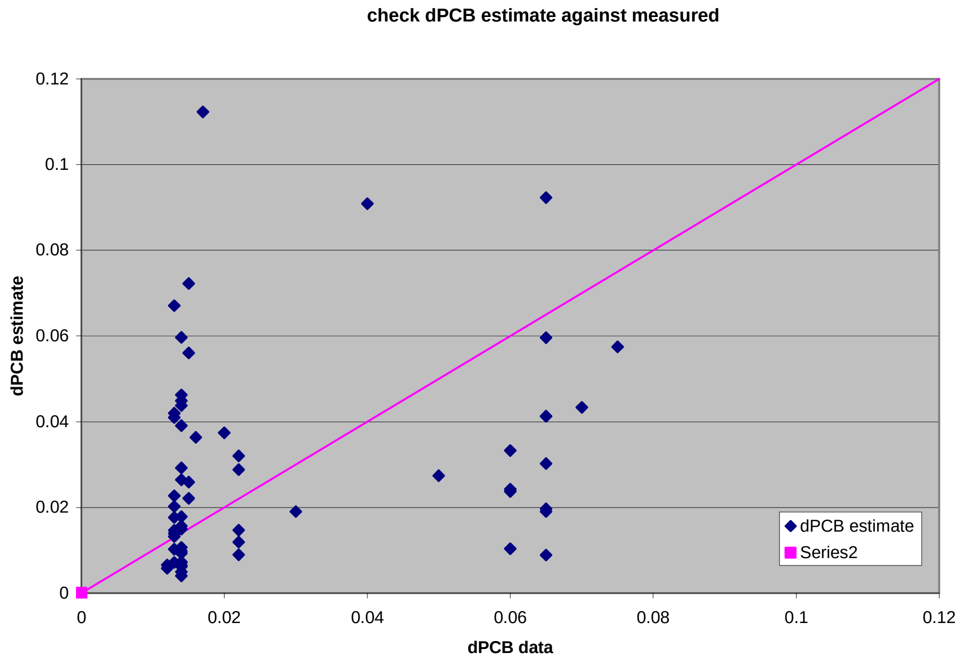


Figure 3

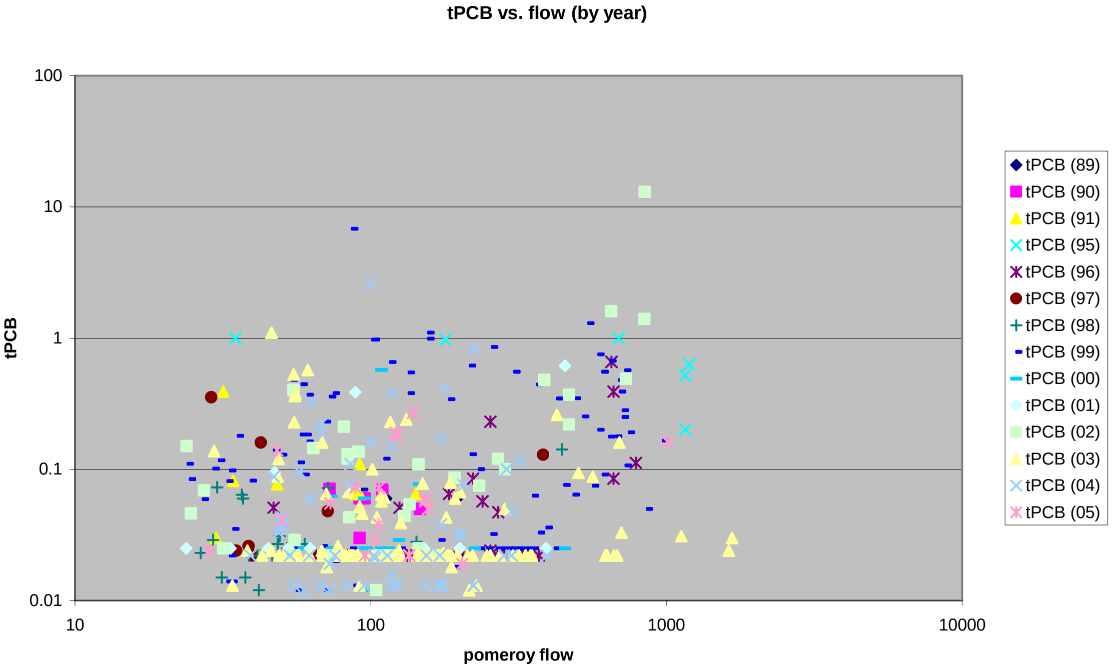


Figure 4

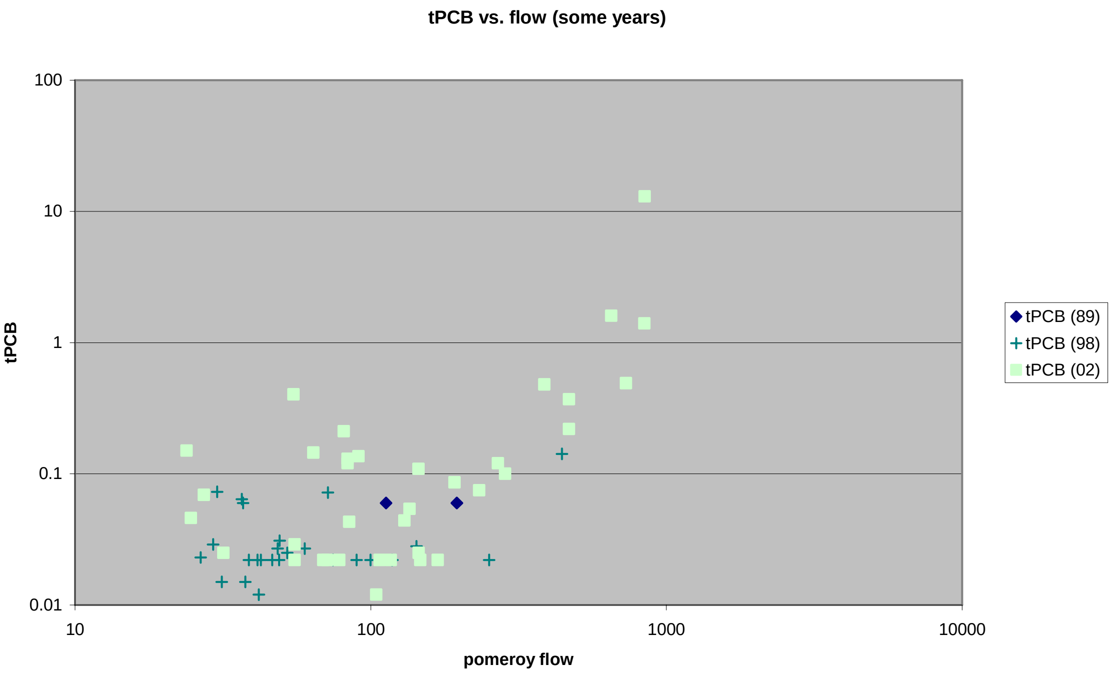
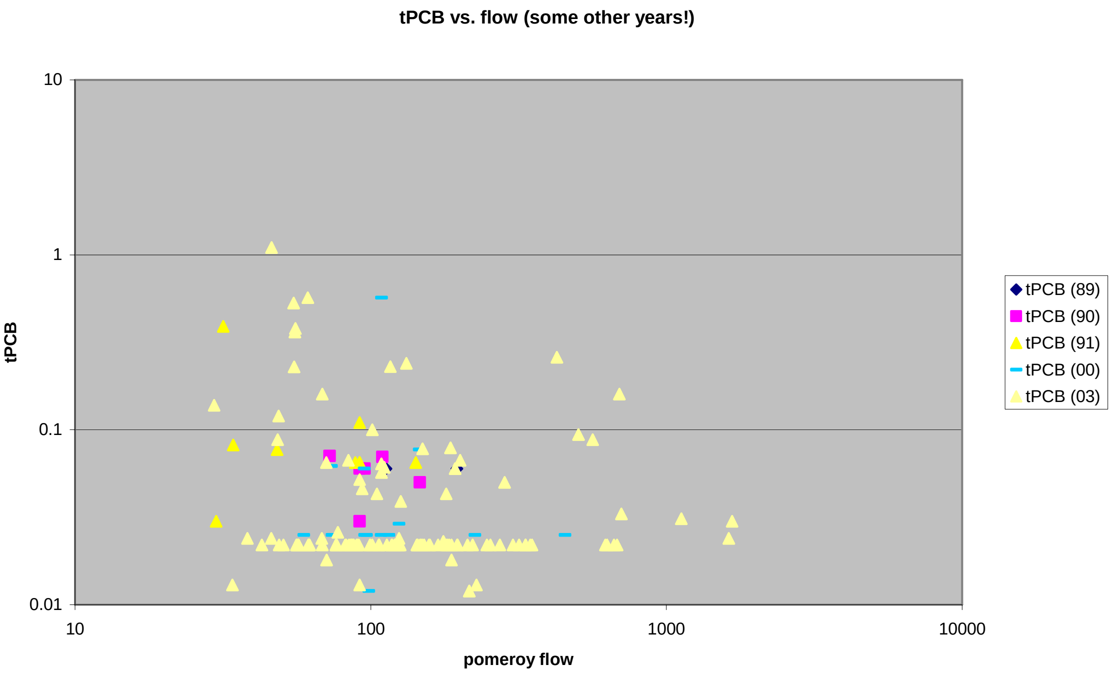
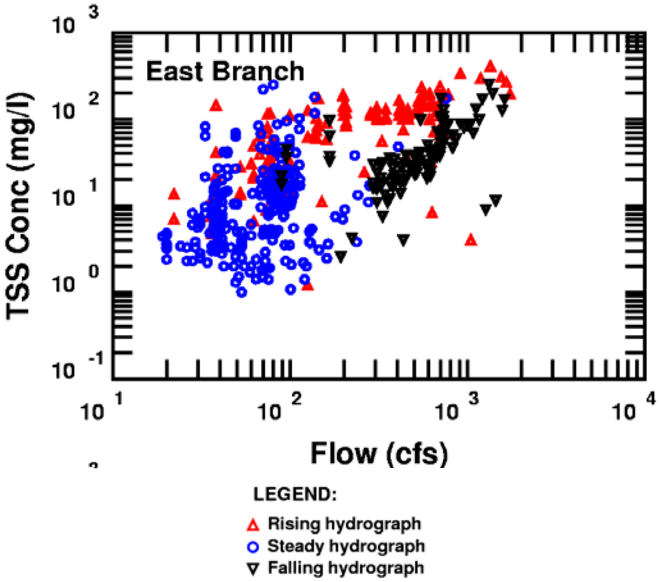


Figure 5



9. I also went back to look at the TSS boundary condition, which has not been changed from the Phase 1 calibration. Figure 6 is a reproduction of the plot displaying the East Branch TSS data as a function of flow (from: Volume 3—Appendix B Hydrodynamic and Sediment/PCB Fate and Transport Model Calibration, Attachment B.3 TSS and PCB Flux Analysis). The TSS data were partitioned according to 3 phases of the hydrograph – rising, steady and falling. TSS data are well-behaved during rising and falling hydrographs, but not at steady (and predominantly low) flow. This makes me wonder again about seasonal factors, intermittent upstream disturbances, or possibly sampling artifacts. It is hard to imagine TSS concentrations varying between 1 and 100 mg/L at base flow. **(6)**

Figure 6



4. Have the sensitivities of the models to the parameterization of the significant state and process variables been adequately characterized?

▶The MVR does a good job of illustrating the sensitivity of the EFDC model predictions to input parameters. Several additional model inputs should be added to the sensitivity analysis, and these are identified below. As pointed out by GE, there are also some sensitivity analysis results that are counter-intuitive; these should be explored and explained.

Sediment transport sensitivity analysis is displayed in terms of peak and mean TSS fluxes, for the entire calibration period as well as 2 different flow events (Figures 5.1-3 thru 5.1-5). It would be useful to also see the sensitivity of the model displayed as changes in bed elevation.

Part of the sensitivity analysis for the sediment transport model should include the sensitivity to the grid resolution because, as discussed previously, the prediction of sediment bed scour may depend on the number of lateral cells in the channel cross-section. This is only needed for a “test reach” of the river channel, preferably the reach where bathymetry transects were measured. Note that Earl Hayter’s previous analysis was not informative on this issue.

PCB fate and transport sensitivity analysis is displayed in terms of peak and mean PCB fluxes, again for the entire calibration period as well as the two flow events (Figures 5.1-6 thru 5.1-11). The sensitivity of the Woods Pond PCB flux predictions at low flow (Figure 5.1-9) is puzzling, since it suggests practically no sensitivity to parameters associated with sediment-water exchange of PCBs. While there may be little exchange within Woods Pond itself, considerable PCB exchange in upstream reaches should be reflected in the sensitivities of these parameters in Woods Pond. Instead, Figure 5.1-9 show that low-flow PCB fluxes in Woods Pond are only mildly sensitive to the upstream boundary inflow, and really nothing else. It remains for EPA to explain how this result is evidence that “the model is properly representing the physics of the river as it relates to the movement of solids and contaminants”.

As stated previously, the initial PCB concentrations specified for the sediment bed should be added as parameters in the sensitivity analysis, for all reaches where they are not independently determined from data. Likewise, the surficial sediment layer thickness should be included as a parameter in the sensitivity analysis. **(5A)**

5. Are the uncertainties in model output(s) acknowledged and described?

▶Considerable effort went into the analysis of model uncertainty in this project, which is perhaps unprecedented for a suite of water quality models of this complexity. The MVR evaluates the uncertainty of the models using approaches based on Monte Carlo analysis. This approach basically assumes that the models are correct, and errors in model predictions arise from uncertainty or variability in parameter values. These are estimated based on subjective expert judgment using probability distributions, and parameters are generally assumed to be uncorrelated. While these assumptions are obviously suspect, they are nonetheless common practice in uncertainty analysis. Analyzing the uncertainty of the linked model predictions is a strength of this work.

The response surface modeling of EFDC uncertainty appeared to be very thorough and comprehensive. EPA has provided a descriptive field for the parameters listed in Table 5.2-8, which hopefully will be incorporated into the report. I would like to see EPA add initial and boundary conditions for PCBs to this analysis, since these are both uncertain factors. I disagree that the RSM approach is impractical or unworkable. The Responsiveness Document (p.2-110 and 111) discusses how uncertainty analysis can be used in context of evaluating differences between remediation scenario predictions. Hopefully, no more than one RSM will be necessary for evaluating uncertainty of these predictions.

The Monte Carlo analysis of the FCM is also quite thorough and informative. I agree with the suggestion to evaluate model uncertainty in terms of 10th and 90th percentiles, as being reasonably conservative, consistent with the distribution properties of the food chain model outputs. This is probably the simplest way to deal with inflation of predictive uncertainty. The critique of “unreasonable” parameter sets producing model predictions outside the bounds of the data, is really a criticism of the subjective estimation of parameter variability. This can be addressed using a Bayesian Monte Carlo approach, which builds posterior (informed) parameter distributions based on computing the likelihood function for each model realization[5]. The method is simple and works very well - although the computational burden of BMC (i.e., effort to achieve convergence in the likelihood function) can be heavy, unless there is some way to substitute the RSM for the numerical model. I don’t know whether this is a practical option for computationally-intensive models such as EFDC.

There has also been much discussion about how residuals from comparisons between model predictions and data can be used to improve the uncertainty analysis. It has been suggested to use residuals to directly characterize uncertainty. This strikes me as a naïve approach, as I would normally consider residual error to be a “best case” estimate of uncertainty. Forecasts will seemingly always be more uncertain than calibration; the question is, how much more? It has also been suggested that reporting average bias (e.g., Tables 4.2-4, 4.2-10 and 4.3-2) masks the uncertainty due to variability in the PCB data itself. The remedy to this is straightforward: calculate and report the variance of model bias. Moreover, statistical model could be developed from the residuals to estimate the distribution of PCB concentrations from mean predictions. This would compliment, instead of replace, the Monte Carlo results. The difficulty I can foresee, is how do you determine the proper assumptions regarding homoscedacity (e.g., is the variance in a constant factor of the mean or simply a constant?). **(5B)**

▶No sensitivity or uncertainty analyses have been presented for the upstream and downstream models. **(8, 9)**

---

[5] Dilks, D.W., Canale, R.P. and P.G. Meier. 1992. Development of Bayesian Monte Carlo techniques for model quality model uncertainty. Ecological Modeling. 62:149-162.

6. Upon review of the model projections of changes in PCB concentrations in environmental media in the example scenarios, are such projections reasonable, using your technical judgment, and are they plausible given the patterns observed in the data?

▶The MVR presented 25-year forecasts made with the linked models, in the Rest-of-River and Downstream domains, for 2 (or 3)[6] hypothetical remediation examples. These were intended to illustrate the performance of the models in forecasting the outcome of remediation alternatives in terms of PCB concentrations in sediment, water and especially fish. These forecast predictions ("projections") are really the unique product of water quality models, and are the only defensible means of anticipating the long-term outcome of remedial action and the relative effectiveness of remedial alternatives. We are asked in this question to address the "reasonableness" and "plausibility" of the projections, which are both subjective criteria. Dom DiToro called this sort of evaluation the "laugh test", where we compare the numerical model's projections to those of the models we carry around in our heads. The modeling team is correct in pointing out that it is not possible to determine whether the example projections are "correct" in terms of accuracy (MVR, p 7-3), because there is really nothing to compare them to except one another.

Model projections for the example scenarios were presented a number of ways in the MVR and its supplements. These included the predicted time series of PCB concentrations in water, at the conventional monitoring locations (Holmes Road, New Lenox Road, …); PCB concentrations in surficial sediment in the mile reaches; and, PCB concentration in largemouth bass in the river reaches 5A thru 6. For EFDC, the projections were also presented as plan maps of the grid displaying PCB concentrations in surficial sediment and soil, and changes in concentration, at the end of the 25 year projections. For the top 6" of the sediment bed, the end-of-projection results are also shown as profiles. Each of these is discussed below.

The projected timeseries of water column TSS and PCB concentrations are plotted together in the figure "Time series of water column concentrations for PCBs and TSS" in *the EPA Response to Questions* document. The base case projection is qualitatively similar to the validation predictions, with TSS and PCB concentrations increasing to high levels during flow events, although over time the base case PCB

---

[6] I am somewhat confused by what is referred to as a "base case" projection in a number of the time series figures (e.g., "Time series of water column concentrations for PCBs and TSS", "Time series of sediment PCB concentrations in each spatial bin" and "Time series of plots of FCM exposure concentrations" in *the EPA Response to Questions* document, and "Food chain model results" in the Document Overview presentation of Example Scenarios, which included a graph subtitled "Base case for the example runs - existing boundary loads"). The "base case" scenario is not mentioned or described in Section 7 of the MVR, or anywhere else I could find in the materials provided for our review. This makes it difficult to evaluate the example scenario projections, because I want to compare them to a "base case" (as is intended by the graphs mentioned above) but am not sure what this is. Initially I thought the "base case" was simply the validation prediction, since the hydrograph and boundary conditions are the same. Upon inspection, however, it appears that the initial conditions for PCBs are somewhat different, most noticeably in Woods Pond. A description of the "base case" scenario should be added to Section 7 of the MVR to resolve this ambiguity. Furthermore, EPA should consider whether a scenario should be added or modified which specifies the East Branch PCB boundary condition to be 20 ng/L, the assumed nominal concentration following cleanup.

concentration projections decline during low-flow periods. PCB concentrations projected for the example 1 scenario are lower than the base case, and projections for the example 2 scenario are lower than for example 1. In each scenario, there is a similar “event responsiveness” of PCB concentrations to fluctuations in the hydrograph. TSS concentration projections are comparable in each scenario, as would be expected since the TSS boundary condition and grain size distribution in the sediment bed are unchanged from the base case. I would consider the water column projections for each of these scenarios to be reasonable and plausible.

The projected timeseries of surficial (top 6”) sediment PCB concentrations in river mile reaches are plotted together in the figure ”Time series of sediment PCB concentrations in each spatial bin” in the *EPA Response to Questions* document. The base case projections for surficial sediment PCB concentrations are qualitatively similar to the validation predictions, and suffer similar errors. PCB concentrations are projected to increase in the first half-mile reach (an apparent artifact), and show almost no change after start-up in the other reaches. The rate of change in PCB concentrations in sediment is (again) unreasonably slow. Even 25 years is not long enough to see a discernable change in surficial sediment PCB concentrations, which I would certainly expect to see. Nothing shown in the MVR or at the peer review meetings convinces me that the model is correct in its predictions of the rate of change in sediment PCB concentrations. For example scenario 1, a decline in sediment PCB concentrations is projected in comparison to the base case, which is more encouraging. The difference is most dramatic in the first half-mile reach (RM 135.13-134.89), and is also apparent in reaches RM129.19-128.88, 128.88-128.69, and 126.99-126.47. In other reaches, especially in Woods Pond and its headwater, there is no difference in PCB concentrations between example 1 and the base case. For example scenario 2, negligible (1 mg/Kg or less; I can’t tell the actual value due to the axis scaling) sediment PCBs are projected for reach 5A due to remediation. In the other reaches, sediment PCB concentrations are projected to be nearly the same or identical to example 1 projections. Although I don’t believe that the model’s projections of rates of change in sediment PCB concentrations are reasonable, at least the relative differences in projected sediment PCB concentrations between the scenarios are plausible.

The grid map diagrams (Figures 7-1 and 7-3) show the cumulative change of PCB concentrations in the top 6” of sediment and floodplain soil over the 25 year projection, for scenarios 1 and 2[7]. It appears that the sediments in reach 5A, which are cleaned up in example scenario 2, recontaminate to 1 or 2 mg/Kg over 25 years, due to continuing transport of PCBs from the West Branch boundary condition. It would be useful to see a similar grid map diagram in which the difference in end-of-projection PCB concentrations was plotted. **(10)**

▶As pointed out by GE, predictions of PCB concentrations in flood plain soils are untested. There appear to be substantial increases in PCB concentrations in the top 6” of much of the floodplain soil in both example projections, which is somewhat surprising. However, since no data have been shown to establish

---

[7] I suppose the comparable graphic for the base case would be the validation figure (Figure 6.2-55) although I am forced to speculate here.

rates of PCB concentration change in floodplain soils, the “reasonableness” of the flood plain projections cannot be determined. Really, this illustrates the absurdity of devoting so many EFDC model grid cells to the flood plain, which it seems dictated using a single footpath of tiles for the grid of the river. **(3C)**

▶The surficial sediment PCB concentration profiles offer a useful alternative view of the cumulative change of PCB concentrations over the 25 year projection. Projections for example 1 show the now-familiar redistribution of sediment PCBs in the high-energy reach 5A, but also significant changes (approaching ± 100% of the IC) in reaches 5B and the upper end of 5C. At the lower end of reach 5C and reach 6, sediment PCB concentrations decline by 20 to 80%. This seemingly contradicts what I interpreted from the “Time series of sediment PCB concentrations in each spatial bin” figures. Could it be that the combination of reach averaging, displayed on a log-axis plotting, makes it difficult or impossible to properly interpret the projection results?

At the Peer Review meeting last month, there were questions and discussion about whether it was reasonable that the EFDC model projected a significant reduction in PCB concentrations in the water column, but little or no change in sediment concentrations. In fact, such behavior is expected in water quality models incorporating both water column and sediment sedimentation. In such models, the concentrations of particle-associated constituents change much more rapidly in water than they do in sediment, whenever there is a change in the constituent loadings or boundary conditions. Since PCB boundary conditions and loadings (actually fluxes from reach 5A sediments) are being manipulated in the example scenarios, the differences in the projections of PCB concentrations in the water column versus sediments are entirely expected and reasonable.

Example scenario projections made by the FCM were presented at the Document Overview meeting (13 - Example Scenarios.pdf) as timeseries of PCB concentrations in largemouth bass. In reach 5A, bass in the base scenario are forecast to have a body burden of 50 mg/kg at the end of the 25 year projection; for example 1, the comparable forecast is 30 mg/kg, and 0 concentration for example 2. In reach 5B, bass in the base scenario are forecast to have the highest PCB body burden at the end of the 25 year projection: 70 mg/kg; for example 1, the comparable forecast is 40 mg/kg, and less than 20 mg/kg for example 2. Further downstream in the Rest-of-River, the remediation examples are projected to considerably less effective in reducing PCB concentrations in comparison to the base case. In reach 5C, bass in the base scenario are forecast to have a body burden of 30 mg/kg after 25 years; for example 1, the comparable forecast is 20 mg/kg, while for example 2 the forecast is 10 mg/kg. At the end of the 25 year projections in reach 6 (Woods Pond), bass in the base scenario are forecast to have a body burden of 50 mg/kg; for example 1, the comparable forecast is 40 mg/kg, while for example 2 the forecast is 30 mg/kg. To me, this progression in the declining effectiveness of remediation as one moves downstream appears reasonable in the context of PCB water and sediment exposure projections for these example scenarios. In “no data” reach 5D, the least reduction in bass PCB concentrations due to remediation is projected. I suppose this may be because this reach is relatively isolated from the rest of the river. **(10)**

▶No scenario projections were presented for either the upstream or downstream model. **(8, 9)**

7. Is the final model framework, as calibrated and validated, adequate to achieve the goal of the modeling study to simulate future conditions 1) in the absence of remediation and 2) for use in evaluating the effectiveness of remedial alternatives?

▶I interpret this to be more than one question, so I will give two answers. If the goal is to forecast PCB concentrations in the Housatonic River in the absence of further remediation, then the current models may be adequate (i.e., "good enough"). In their current form, the EPA models (and probably many much simpler models) can provide meaningful answers to many questions about the future of PCBs the Housatonic River. The model calibration and validation show the relative magnitude of PCB inputs to system, and the resulting predictions and diagnostics are instructive in terms of evaluating remediation targets. For example, eroding banks and sediment deposits in reach 5A are clearly targets for remediation. The 25-year validation and example scenario projections also address the need and possible effectiveness of remediation in the Housatonic River. Following the remediation of the East Branch reaches, upstream PCB concentrations of 20 ng/L and 0.4 mg/kg expected; these concentrations alone may support gamefish concentrations of 1 ppm or more in the PSA. In the absence of further remediation, PCB concentrations will decline very slowly from the current levels. Although water-column PCB concentrations will decline downstream of remediation conducted at the reach scale, surficial sediment PCB concentrations will be relatively unaffected in downstream reaches. The predicted response of PCB concentrations in fish to remediation is expected to be somewhat similar to the sediment. To substantially reduce (e.g., >50%) PCB concentrations in Reach 5D and 6 gamefish over 25 years appears to require remediation well beyond the spatial extent of hypothetical example 2. None of these conclusions require a highly accurate model.

However, if the goal is to forecast PCB concentrations to evaluate the effectiveness of further remediation, and to distinguish between the effectiveness of various remedial alternatives, then I believe that the current models are not adequate. Although they are good, they are not as good as they could be. I assume that GE will not willingly undertake additional remediation activity in the Housatonic River, unless EPA can convincingly demonstrate that such remediation is necessary and that the remedy will be effective. At their current stage of development, the EPA models are not ready to make such a demonstration. I am arguing that further model development and testing is warranted, and that EPA should be motivated to do so because the financial stakes are high. Otherwise, GE will likely exploit weaknesses in the models to argue against the need for further remediation. **(12)**

▶Outlined below are the steps I believe are necessary to make the EPA models the best possible tools to accomplish the goals of the modeling study for the Housatonic River:

- Revise the MVR to incorporate supplemental material (*EPA Response to Questions* document , Document Overview presentations, etc.) and remove provisional results (mass balance diagrams, downstream model w/o boulders, …) that have been superceded since the MVR was released.

- Make a number of near-term corrections to EFDC, which can be addressed/resolved within a ~1 yr time frame)
    - Correct parameterization errors identified in EFDC model:

        Reduce thickness of surficial sediment bed layers;

        Parameterize vertical mixing rates as functions of benthos density & vertical position in sediment bed;

        Increase diffusive PCB fluxes by calibration and parameterize spatially as functions of benthos density;

        Reanalyze SEDFLUME data using resuspension parameters constrained by literature (e.g., shear stress exponent n=2.6±0.3) and recalibrate deposition rates. Lick, Ziegler and coworkers have published much guidance on the specifics of parameterizing sediment resuspension[8], some of which the modeling team has chosen to ignore;
    - Revise TSS and PCB boundary conditions to correct bias evident in low flow range.
    - Test sensitivity of EFDC hydrodynamics, sediment transport and PCB transport simulations to alternative grid resolutions in a "test reach" of the PSA. In all prior SEDZL applications I am aware of, at least a 2-dimensional model was used for hydrodynamics and sediment transport. In river systems, this was usually a vertically-integrated model. Furthermore, on rivers that bend as much as the Housatonic, a curvilinear grid has commonly been applied. Since the EPA modeling team has elected not to follow these conventions, they should at least demonstrate via numerical testing that their primarily 1-dimensional model of the Housatonic River channel produces comparable results for sediment and PCB transport.
    - Investigate methods to economize on hydrodynamic and sediment transport simulations (e.g., using steady state design storms and flow-duration statistics, such as the Saginaw River SEDZL application[9])
    - Consider using the calibrated and validated EFDC model to build a simpler box model of the river, if the two models can be shown to produce comparable results. The simpler model could be particularly useful for forecasting uncertainty. Don Mackay suggested this "second simplified model" approach for modeling hydrophobic contaminants in the Niagara River[10].

---

[8] See, for example: A Quantitative Framework for Evaluating Contaminated Sediment Sites. SETAC 20th Annual Meeting, Philadelphia, PA, November 14-18, 1999.

[9] Cardenas, M. and W. Lick. 1996. Modeling the Transport of Sediments and Hydrophobic Contaminants in the Lower Saginaw River. *J. Great Lakes Res.*, 22(3):669-682**.**

[10] McLachlan, M. and Mackay, D., "A Model of Contaminant Fate in the Niagara River", report prepared for Environment Canada (1987).

- Develop and implement a process of ongoing model validation, using data collected as remediation progresses. It goes with out saying that modelers want to continue modeling… However, the reality is that no model is truly ever “finished”. So long as new data are collected, model refinement must be expected and accommodated by managers and decision makers. **(13)**

▶I do not believe that either the upstream or downstream models are sufficiently developed to be adequate to address the goals of the modeling study. Their development appears to have been rushed, an impression reinforced by the fact that they apparently did not exist a year ago. Sensitivity and uncertainty of the upstream and downstream models have not been reported, and documentation of the development of these models has been inadequate. In its present status, the upstream model is incomplete and not fully validated. PCB fate and transport have not been simulated, and insufficient model-data comparisons have been made and/or reported. Incorporation of the upstream model in the MVR appears premature, given that substantial additional development efforts will be required before it can be used reliably. Likewise, the downstream model has not been sufficiently validated. Limited model-data comparisons suggest there may be significant bias in PCB concentration predictions. There also appears to be a lack of adequate and appropriate data for this significant extension of the model domain. In their current states of development, neither the upstream nor downstream models are suitable for use in modeling future conditions in the Corrective Measures Study. I am hopeful that the inadequacies of the upstream and downstream models can be addressed through continuing monitoring and modeling activities. **(8, 9)**

APPENDIX A.4

COMMENTS OF MARCELO H. GARCIA

FINAL COMMENTS ON HOUSATONIC RIVER MODEL VALIDATION REPORT

By Marcelo H. Garcia, PhD

According to the charge for the Model Validation Report Peer Review, the goal of the modeling study is to develop a tool that will:

- Predict future concentrations in various media (e.g., sediment, fish, and water);
- Assess relative performance among remedial alternatives against baseline conditions; and,
- Be the best estimate available of the potential magnitude of the expected reductions in exposure and, thereby, provide useful information in evaluating the performance of remedial alternatives.

<u>Question 1:</u>

Considering the changes implemented in the Phase 2 Calibration, does the model as calibrated and validated, based on your technical judgment, reasonably account for the relevant processes affecting PCB fate, transport, and bioaccumulation in the Housatonic River to a degree consistent with achieving the goal of the modeling study?

<u>Response:</u>

▶While a substantial effort has gone into improving the capabilities of the Housatonic River Model since its calibration, it is clear that there is still a long way to go before the model can truly be used as a “predictive tool” to quantify future spatial and temporal distributions of PCBs (both dissolved and particulate forms) within the water column and the bed sediment.

This reviewer is of the opinion that the Housatonic River Model will need to be continuously improved until all the processes (biological and physical bed sediment mixing, streambank erosion, floodplain deposition, etc.) relevant to the transport and fate of PCBs are not only accounted for in the model but are also well represented and based on sound knowledge of sediment transport mechanics, stream biology, ecology and morphodynamics. **(1, 3C)**

▶It is not enough to state that a given process is accounted for when the representation of the process is deficient and its implementation in the model is done with algorithms and assumptions that are not based on the physics of the process and cannot be supported by field observations. For instance, streambank erosion has been recognized as a very important process since large amounts of PCB-laden sediments can enter the river during medium to high flows. However, fluvial erosion and mass failure of the stream banks are currently modeled simply as functions of flow discharge, a rather crude approximation, which severely limits the capabilities of the model to assess the effect of this process.

How can one determine what regions of the banks are more prone to erosion and, therefore, need protection if the model does not compute the distribution of flow velocity and bed shear stress along the banks? Could native vegetation or other bioengineering techniques be used to protect the banks of the Housatonic River against erosion or more hard-core solutions such as rip-rap are needed? Is it possible to use a combination of both? As it stands now, the model can not be used to compare the merits of such remedial alternatives, which is one of the goals of the modeling study as stated above. **(1D, 12)**

▶Arguably the main shortcoming of the model is not the Housatonic River model itself but rather its numerical implementation which, in my opinion, has rendered the model calibration and validation unreasonably difficult. Assuming that all the processes are eventually accounted for in the model and the right algorithms are used to simulate them, the most vexing issue will continue to be the characteristics of the computational grid used to simulate the river hydrodynamics, sediment transport and associated PCB transport and fate. The computational mesh (Figure 3-6) is “hard-wired” to water stages in the floodplain associated with a flood having a recurrence time of 10 years. This results in the main river channel being modeled with the same spatial resolution (i.e. a single computational cell) regardless of the flow discharge, not just the 10 year flood. **(3A)**

▶With the current computational mesh configuration, the application of the model is computationally taxing, resulting in extremely long execution times. This has made the calibration and validation of the model a very difficult exercise and will most likely affect the analysis of remediation alternatives in the future. **(4)** ▶At the same time, the use of a course grid hinders the ability of the model to resolve the flow velocity distribution inside the main channel as well as the associated boundary shear stresses throughout its wetted perimeter. Given that erosion and sediment transport vary non-linearly with increments in flow velocity and shear stress, it is rather imperative to find a way to increase the spatial resolution of the model so that at least three “streamtubes” are used to model flow and sediment transport in the main channel of the Housatonic River. This will in turn facilitate and make more meaningful the computation of streambank erosion as well bed sediment erosion and resuspension. **(1D, 3A)**

▶While the model capabilities have been improved since its calibration, my professional opinion is that the validation of the model is not complete at this stage. In particular the model does not capture the variability and spectrum of PCBs concentrations observed in river. There are several aspects of the model that need to be improved as stated above before validation of the model can be accomplished. **(2B)**

▶Due to the nature of the problem and the scale of the river system under consideration, it is unlikely that a full validation of the model will be possible in the near future. From a practical point of view, however, the main question in my mind would be: is the model good enough to capture most of the process responsible for the transport and fate of PCBs so that it can be used to assess the merits and drawbacks of different remediation strategies in the Housatonic River?

Right now my answer would be no but the effort by EPA and its consultants has provided a good foundation towards the goal of having a useful tool that can be used to help the Housatonic and other rivers experiencing similar problems. **(12)**

<u>Question 2</u>:
Are the comparisons of the model predictions with data sufficient to evaluate the capability of the model on the spatial and temporal scales of the final calibration and validation?

<u>Response</u>:

▶Of particular relevance for this reviewer is the validation of flood plain deposition throughout the river system as well as sediment erosion and deposition in Woods Pond, since the record shows that this pond is a major deposit for PCB-contaminated sediments.

Regarding the sedimentation of Woods Pond, the calibration/validation exercise has produced results showing rates of (*net*) sedimentation for the period going from 1979 to 2004 (Figure 6.2-34). The results seem reasonable with predicted sedimentation rates (mm/year) that are within an order of magnitude of sedimentation rates determined from Cesium data. The model also reproduces fairly well the distribution of PCBs with depth as observed from sediment cores. However, as stated below, the model shows some bias when predicting temporal variations of surface PCB concentrations in Woods Pond.

Woods Pond is perhaps one of the few locations in the Housatonic River where the capabilities of the model to predict the spatial and temporal distribution of sediment and PCB can be extensively tested due to the large amount of observations available. This has been done to some extent but more effort should go into this since the pond could become a major source of PCBs to the downstream portion of the river during a major flood event. **(2A, 2C)**

▶Regarding floodplain sedimentation, the model has been used to predict process-based (advection, erosion, deposition, volatilization, etc) sediment and PCB fluxes (Kg/year) for the main channel and the floodplain, over the validation period (Figures 6.2-62 and 6.2-63). The analysis results in an estimate of yearly fluxes. However, it is well known that sediment erosion and transport is most prominent during flood events, which can have duration of a few hours to several days. Thus it would seem that the model should be validated for time scales that are relevant to the processes involved.

Since a large percentage of the river system consists of floodplain, it is important to ensure that the model can indeed capture the process of floodplain sediment/PCB deposition. To this end, a simple one-dimensional approach was suggested that could be used to estimate a “Floodplain Dimensionless Number” for different reaches of the Housatonic River (Garcia, 2006-technical note on floodplain sedimentation). Once such number is calibrated for each reach, it will be possible to test if indeed the model can predict floodplain depositional rates that result in similar Floodplain numbers for different flooding conditions (hydrologic event scale), thus helping in the overall validation of the model. This approach will clearly show the capabilities of the model. If there is any hope of predicting PCBs fate in the floodplain, sedimentation has to be properly simulated by the model. **(3C)**

▶As mentioned in the response to question one, the size of the computational grid makes it difficult (and meaningless) to compare model predictions for processes that take place at hydraulic and sedimentation scales determined by the flow rate and the main channel width. This is not the case for Woods Pond and the floodplains were the size of the computational grid is appropriate for the scale of the flow field. **(3A, 3B)**

<u>Question 3</u>*:*
Is there evidence of bias in the models, as indicated by the distribution of residuals of model/data comparisons?

<u>Response</u>*:*

▶As shown in MVR Figure 6.2-50, the model predictions of variation with of surface PCBs concentrations are not consistent with the field observations in Woods Pond. The model seems to underpredict the observed values in the time period from 1979-1984 and to overpredict the observations in the time period from 1995 to 2000.

There is evidence that model predictions do not match well with observed surface sediment concentrations of PCBs in reaches 5A, 5B, and 5C as displayed in Figure 6.2-49. **(2A)**

▶Model predictions in Figures 6.2-45 and 6.2-46, show that during low flow conditions the model does not capture observed longitudinal gradients in the water column. This results in the model underpredicting PCB fluxes from the sediments along Reach 5A, which in turn could lead to an under estimate of the bed sediments PCB contributions to fish in the Food Chain Model. **(1B)**

<u>Question 4</u>:
Have the sensitivities of the models to the parameterization of the significant state and process variables been adequately characterized?

<u>Response</u>:
▶Given the facts that several processes are yet to be fully characterized in the model and that a coarse computational grid has been used for the model

calibration, I do not think that a full sensitivity analysis can be conducted at this time. **(1, 3A, 5A)**

▶In the case of low flow conditions, the model predictions seem to be very sensitive to diffusion parameters, suggesting the representation of pore water diffusion as well as the thickness of the so-called “mixed layer’ need to be carefully analyzed. **(1B, 3D)**

Question 5:
Are the uncertainties in model output(s) acknowledged and described?

Response:
▶While uncertainty in model inputs and outputs was recognized, the nature of the model does not allow for a conventional uncertainty analysis. Once all the processes are accounted for and well represented in the model and assuming that the EFDC code can be run more efficiently, it might be feasible to conduct an uncertainty analysis. **(4, 5B)**

Question 6:
Upon review of the model projections of changes in PCB concentrations in environmental media in the example scenarios, are such projections reasonable, using your technical judgment, and are they plausible given the patterns observed in the data?

Response:
▶Model projections seem reasonable but I would have liked to see more potential scenarios, particularly for extreme conditions such as low flow summer-like conditions and floods. **(10)**

Question 7:
Is the final model framework, as calibrated and validated, adequate to achieve the goal of the modeling study to simulate future conditions 1) in the absence of remediation and 2) for use in evaluating the effectiveness of remedial alternatives?

Response:
▶I do not think that the model is ready to accomplish the goals of the modeling study. If the model is to be used to simulate future conditions, first is has to adequately simulate the existing conditions throughout the river and its floodplain. **(13)**

Overall, this is a very challenging undertaking but the study and modeling of the Housatonic River is a worthwhile effort that will hopefully benefit future generations.

APPENDIX A.5

COMMENTS OF FRANK GOBAS

Final Response to Charge Questions

Housatonic River

Frank Gobas, Ph.D., Professor

School of Resource & Environmental Management

Simon Fraser University

July 20, 2006

Response to Charge Questions

Charge Question 1

"Considering the changes implemented in the Phase 2 calibration, does the model as calibrated and validated, based on your technical judgment, reasonably account for the relevant processes affecting PCB fate, transport and bioaccumulation in the Housatonic River to a degree consistent with achieving the goal of the modeling study."

Since the goal of the modeling study is to address 6 model goals (Table 1), I have organized my response to charge question #1 by specifying the various goals of the model.

Re. Model Objectives #1, #5 and #6 on the issue of temporal response:

The main goal of the model is to is estimate the temporal response of the PCB concentrations in the River as a function of various remediation strategies and natural recovery. The model's temporal response is key to (i) the quantification of future spatial and temporal distribution of PCBs (model objective #1), both dissolved and particulate forms) within the water column and bed sediment; (ii) the estimation of the time required for PCB-laden sediment to be effectively sequestered by the deposition of uncontaminated material (model objective #5) and (iii) the estimation of the time required for PCB concentrations in fish tissue to be reduced to levels established during the risk assessment process, that no longer pose either a human health or ecological risk (model objective #6) as well some of the other model objectives.

▶The temporal response of the model is based (i) on the parameterization of the sediment-water exchange rate of PCBs (which is the rate determining step in the depuration of the River) and (ii) on model calibration. With regards to the parameterization of the sediment-water exchange of PCBs, there are several model parameters including the depth(s) of "accessible" and "non-accessible" bottom sediments, diffusion rates, bioturbation rates and subduction velocities that are currently difficult to measure or assess. The best strategy for parameterization is to use the best empirical data and best expert judgment possible. I am not convinced that this was achieved in this study. Dr. Lick has provided several papers to demonstrate that the current parameterization of the sediment-water exchange process in the model is not consistent with some key studies and observations. Also, I place considerable weight on Dr. Lick's expert judgment and his lack of confidence in the selection of parameter values. However, that said, it is unknown at this time whether an alternative parameterization of the sediment-water exchange processes will produce a significantly different and better characterization of the temporal response of the PCB concentration in the River. **(1)**
▶The model calibration has been complicated by the lack of a significant change in PCB concentration data in the River during the calibration period. Hence, the model calibration sheds little light on the issue of whether the model is able to account for

changes in PCB concentrations over time that may occur as a result of remediation. The main conclusion from the calibration is that the temporal response of PCB concentrations in the River is likely slow, but it cannot tell us accurately how slow.

Evidence to indicate that the current temporal response of the model may be flawed was presented by Dr. Connolly, who suggested that the half-life time of PCBs in the River is likely is closer to 36.5 yr (i.e. 0.693/0.019) than the current model prediction of 110 yr (i.e. 0.693/0.006). I think that Dr. Connolly is probably correct, but the uncertainty in the estimates of the half-lives is large and I have not been convinced that the current PCB concentration data in the River allow one to distinguish between these two estimates.

In conclusion, the current characterization of the key temporal aspects of the model is too uncertain to make reliable long term predictions of the concentrations of PCB in water, sediments and biota of the Housatonic River. This said, it is clear from the empirical data and the model calculations that the long term temporal response to changes in PCB loadings is slow. There appears to be very large standing mass of PCBs in the River sediments and its floodplains and the sediment removal rate from the PSA is relatively small. Natural recovery therefore can be expected to take a long time, with a system half-life time in the order of decades. Given this slow response time, the calculation of a more accurate natural recovery time may <u>in some cases</u> be inconsequential. Under certain conditions, the model, as it stands, may already be sufficient to address the long term temporal response of PCB concentrations in the water column and bed sediments (model objective #1).

The reason for the model’s ability to convincingly assess the long term temporal response of the PCB concentrations in water, sediments and biota of the River are twofold. First, no significant long-term-changes in PCB concentrations over time were observed in water, sediment or biota throughout the PSA during the calibration period. Secondly, the key parameters controlling the temporal response of PCB concentrations are difficult to measure or estimate. **(2A)** Based on this, the following recommendations can be made:

Recommendations:

▶The modelers should explore alternate parameterization schemes (possibly with the help of Dr. Lick) of the sediment-water exchange processes with the goal to select model parameterization schemes that are defensible based on available laboratory and field observations. **(1)**

▶The modelers should revisit the calibration of the temporal response of the model based on a more in-depth statistical analysis of the available model calibration data. I specifically refer to the sediment concentration data as a function of time. Model calibration schemes should be evaluated in terms of their ability to reproduce statistically validated temporal PCB concentration data. **(2A)**

▶The modelers should improve the calibration of the model by lengthening the model calibration period. This most likely involves a continuation of monitoring the response of PCB concentrations in the River to the remediation efforts that have recently taken place. Over time, a more definite change in concentration may take place, which can be used to better calibrate the model. **(13)**

Re. Model Objective #1 on the issue of Space:

▶The model report reveals insufficient information to determine whether the capability of the model to make accurate predictions of spatial differences in PCB concentrations is adequate. The reason is that the calibration of the spatial characteristics of the model revealed that differences in PCB concentrations (e.g. see Fig 6-2-45 and 6-2-46 Model Validation Report) among River sections are small while variability in the PCB concentrations in the sediments at individual river locations is high. **(2B)**

▶The spatial resolution of the environmental fate and food-web model is unnecessarily complex. **(3A, 4, 11)** ▶A much simpler spatial model is consistent with the empirical data and the current model calculations. The hydrodynamic model does not run on a low spatial resolution. However, it is possible to run the hydrodynamic model on its optimal spatial resolution and aggregate the hydrodynamic data for a much simpler PCB environmental fate model. The latter would also reduce the model run-time from the current, unacceptable 30 to 50 d. This would also make the model more transparent due to greater simplicity. The modelers could argue that in terms of predicting spatially varying concentrations the current multi-compartment model is consistent with a much simpler lower spatial resolution model, so there is no need to develop the simpler model. This is true for the model's current application. However, the modelers should keep in mind that when the current model is applied to make predictions of the impact of remedial actions on PCB concentrations as a function of space (i.e. model objective #1), the model may predict spatial differences in concentrations that have not been "validated" or "ground truthed". **(3A, 4)**

Recommendations:

▶Better spatial statistical methods should be used to explore the current spatial PCB concentration data for spatial trends in the available PCB concentration data. **(2B)**

▶The modelers should consider developing a simpler spatial representation of the environmental fate and food-chain model by decoupling the hydrodynamic and sediment transport model from the environmental fate and food-chain model. The current high spatial resolution of the model is unnecessary for the environmental fate and food-chain model. **(3A, 4)** ▶Also, the time-step requirement for the hydrodynamic model is too onerous and unnecessary for the environmental fate and food-chain models. **(3A, 4, 11)** ▶This change in model design will make it possible to make many model runs within a reasonable computational time. **(3A, 4)**

Re. Model Objective #2:

▶With regards to achieving model objective #2, i.e. ability to quantify the historical and current relative contributions of various PCB sources to PCB concentrations in water and bed sediment, I do not think that the model reasonably accounts for the relevant fate processes. The reason is that the model has difficulty assessing the amount of historical PCB mass that is "accessible" by the River. Difficulties in the selection or determination of an "accessible" sediment layer and decisions regarding the inclusion or exclusion of flood plains as accessible sources of PCBs to the Housatonic River contribute to the overall difficulty in assessing the current mass of PCBs in the River. As a result it is difficult to assess the relative contribution of any current inputs of PCBs into the River. The lack of any significant change in the PCB concentrations in the River over time over the period that significant removal of PCB sources in the immediate vicinity of the GE facility remediation took place may be an indication that historical sources of PCBs throughout the PSA are likely the main contributor to current PCB concentrations in the River. **(1)**

Recommendation:

▶An alternative to the difficult characterization of the current mass of PCBs in the River, is using the model to investigate under which set of model parameters historical sources are the main contributor to the PCB concentrations in the Housatonic River and judge whether these model parameter sets are reasonable. This would avoid having to characterize the actual current mass of PCBs in the River. **(13)**

Re. Model Objective #3:

▶Due to the heavy reliance on calibration during model development, the relative PCB concentration data in water, sediments and biota that are calculated by the model are overall consistent with the model predictions of the mean concentrations. **(11, 13)** ▶The model can therefore be used with confidence to address the relative contributions of <u>current</u> PCB sources to bioaccumulation in target species (i.e. model objective #3). Dr. Connolly argued that the potential underestimation of the overall PCB depuration rate in the River affects estimates of the relative contributions of current PCB sources to bioaccumulation in target species. While this is correct, I do not think that this will have a significant effect on the derivation of current relative sources of PCBs to fish because PCB concentrations did not vary significantly over the time period that calibration was performed.

The application of the model of the model to quantify the <u>historic</u> contributions of various PCB sources to bioaccumulation in target species is dependent on the time dependent capabilities of the environmental fate model, which are more uncertain (see discussions above). **(13)** ▶The FCM model can be expected to properly estimate the relevant contributions of water and sediment concentrations as sources of PCB bioaccumulation in benthos and fish species from PCB concentrations in water and sediments delivered by the EFDC model. **(11, 13)**

▶Re. Model Objective #4:

The report indicates that the model reasonably accounts for the role of storm events. The model appears to do a good job in describing the short term temporal response of river flow rates in response to storm events. The comparison of simulated and observed river flows is more than adequate. However, the excellent flow predictions do not translate in equally impressive TSS concentrations. Figures 6.2-19 to 6.2-22 show considerable discrepancies between observed and predicted short term TSS concentrations at Holmes Road, Lennox Road and Woods Pond Outlet. On the other hand, Figures 6.2-39 to 6.2-41 show that the short term PCB concentration variations in the water phase (assuming that the concentration data displayed are water column tPCB concentrations) are reasonably well predicted. The reason for the poor predictability of TSS concentrations, but good predictability of the PCB concentrations in the water column is somewhat surprising and unclear. However, the data presented produce confidence in the model's ability to make reasonable predictions of the effects of storm event(s) on the redistribution of PCB-laden sediment in the study area (model objective #4).

Recommendation:

I recommend that the modelers investigate the source of error in the estimation of the TSS. **(2A)**

Charge Question #2

Is there evidence of bias in the model, as indicated by the distribution of residuals of model/data comparisons?

There is evidence of bias in certain model outcomes in the model validation.

▶First, in the majority of comparisons between measured and model predicted TSS concentrations, the TSS concentrations are over predicted. This does not appear to have a corresponding impact on the calculation of the PCB concentrations in water. The effect of this bias on the calculation of PCB concentrations under current conditions appears to be low. However, the bias may become important when the model is applied under different conditions. It is therefore important to explore the reasons for the bias in TSS and make appropriate corrections. **(6)**

▶Secondly, the combined or linked model (Table 6-4.7) shows a systematic underprediction of the mean PCB concentrations as high as a factor of about 2 for some of the species. Overall, including all species, the underprediction of PCB concentrations in biota is about a factor of 1/0.60 or 1.67. This systematic bias does not appear to be due to the bioaccumulation model itself. Judging from Table 6-4.6 (Model Validation Report), the bioaccumulation model itself appears to have little or no systematic bias. Judging from Figures 6.3.3 to 6.3.8, the uncertainty in the characterization of the model bias is large. The authors could have calculated the standard deviation of the mean, but did not do this. If they would have done this, they would have found that the standard deviations of the model bias for any of the PCB concentration in fish data sets are quite substantial due to the large variability in the observed PCB concentrations in the biota. This means that while the mean model bias for the PCB concentration predictions in fish is relatively low, the uncertainty of the mean model bias is high. I think that it is important in any model to be upfront about the ability of the model to make predictions of reality. The reporting of the model bias without its uncertainty is misleading in my view. The reality is that PCB concentrations in biota vary substantially and we do not really understand why this is. So, we should not pretend that we can predict PCB concentrations in fish with the accuracy that the mean model bias measures suggest. I therefore suggest that the authors provide a full reporting of the model bias of the PCB concentrations (i.e. report uncertainty in the mean model bias) using the linked model and interpret the findings in terms of model uncertainty when the model is applied. **(11)**

▶Thirdly, the analysis presented by Dr. Connolly during the June 28 Public Meeting indicates that there may be a bias in the temporal response of the model, i.e. it appears that the rate of temporal response of PCB concentrations in the River is underestimated by the model. However, there is considerable uncertainty in the data on which the calculations are based and it is unclear whether the uncertainty in the data is sufficiently small to distinguish between the temporal response rates calculated by the model and Dr. Connolly’s analysis. A bias in the temporal response of the model could have large implications for model projections of remedial actions. More detailed evaluations of the bias in temporal response need to be conducted. This should involve the calculation of PCB concentration response times from observed data and the comparison of the measured response times to the calculated response times to determine whether a bias exists. Alternative model parameterization schemes may need to be explored to investigate whether the model bias (if it exists) can be reduced. **(2A)**

Recommendations:

▶I recommend that the modelers explore model parameterization schemes that remove bias in the TSS model predictions. **(6)**

▶I recommend that the authors calculate the standard deviations of the mean model bias of the linked model and make appropriate corrections to account for the systematic underprediction of the mean PCB concentrations in the biota by the model. The systematic underprediction of the mean PCB concentrations in fish by the model needs to be either corrected or recognized when the model is applied. **(11)**

▶The possible bias in the model’s temporal response of the PCB concentrations in the River needs to be investigated by comparing observed and predicted PCB concentration decline rates over time. In case, significant bias exists, alternative model parameterization schemes need to be explored to improve the long term temporal response of the model. **(2A)**

Charge Question #3

Are the comparisons of model predictions with data sufficient to evaluate the capability of the model on the spatial and temporal scales of the final calibration and validation

▶With regard to the temporal scale, the comparisons of model predictions with data are not sufficient to evaluate the capability of the model to make accurate estimates of the temporal response of PCB concentrations in the River. As discussed under charge question #1, this is largely due to the considerable variability in observed PCB concentrations and the lack of a significant decline in PCB concentrations over the calibration period. Hence, a temporal trend is difficult to discern from the data and the lack of temporal trends makes it very difficult to evaluate the temporal characteristics of the model. With regards to the spatial scale of the model, a similar conclusion can be reached at. The small-scale spatial variability in the PCB concentration in the sediments is so great that spatial differences in PCB concentrations among River sections are difficult to discern. The comparison between observed and predicted concentrations therefore provides little information with regards to the capability of the model to make accurate predictions of PCB concentrations as a function of time or space.

Recommendations:

Dr. Connolly's analysis suggests that with the application of suitable statistics it may be possible to use the current data sets to better characterize temporal and spatial PCB concentration trends. Spatial statistics and the application of geographical averaging methods may help to better characterize spatial trends in the data that can be used to evaluate the applicability of the model. Temporal trend analysis can be used to discern temporal trends. I recommend that this is done as it may provide better data for a comparison of observed and predicted concentrations. **(2A)**

▶A second suggestion is to lengthen the model's calibration period. The calibration period for the current model was too short for an evaluation of model capability. This would involve the continuation of monitoring programs with the objective to develop PCB concentration data over a longer period of time. **(13)**

Charge question #4

Have the sensitivities of the models to the parameterization of the significant state and process variables been adequately characterized?

▶The modelers have done a good job describing parameter sensitivities. The latter is not easy as there are many parameters and the sensitivities of each of them depend on the values chosen for the others.

However, there is one area in the model where conducting a thorough sensitivity analysis is crucial because of the lack of good calibration data and difficulties in model parameterization. This is in the component of the Environmental fate model that controls the long term temporal response of the model. The selection of the thickness of sediment layers, diffusion rates, and bioturbation rates are all very difficult. Hence, a thorough sensitivity analysis is crucial to determine the bounds within which temporal changes in concentrations can be expected to occur. This analysis is not presented in the current reports and supports my conclusion that the long term temporal response of the model is too uncertain to make meaningful predictions.

Recommendations:

Given the lack of calibration data, the sensitivity analysis is probably one of the few things that can be done to increase confidence in any estimates of the temporal response in PCB concentrations calculated by the model. I recommend that it is considered and added. For this sensitivity analysis to be doable and convincing, the modelers could just focus on the part of the model that controls the long term temporal response of the model. **(5A)**

Charge question #5

Are the uncertainties in the model output(s) acknowledged and described?

▶1. While the report presents several efforts to calculate uncertainty, the model uncertainties are not fully acknowledged. In essence, the report presumes that the model’s ability to calculate mean concentrations is sufficient to address the goals of the study. I think that this is a flaw in the study design because the goal of the model is to compare PCB concentrations resulting from different remedial scenarios. Such an application involves the comparison of mean concentrations. However, comparing mean concentrations alone is insufficient to determine the significance of the differences in mean concentrations. The calculation of the statistical significance of a difference in the means is required. The latter is a well established practice in scientific and engineering studies. I do not think that a convincing rationale is presented for why this practice is not applicable in this study. The authors argued that ecological receptors (including fish), due to their continuous movement, tend to be exposed to a large variation in concentrations, which get “averaged out” to produce an internal concentration in the fish that corresponds to the average or mean exposure concentration rather than the variation in concentrations (l. 10-14, p.15 Responsiveness document). Figures 6-2-35, 6-2-56, 6-3-3 to 6-3-8 (Model Validation Report), and similar figures in the calibration documents demonstrate a one-to two order variability in predicted concentrations of PCBs, which does not differ substantially between fish and sediments (i.e. the variability in fish concentrations does not appear to be any less than that in suspended solids). This indicates that the original assumption that PCB concentrations in fish may be less variable than the sediment concentrations may not hold.

The model's capability to estimate only mean concentrations, makes it difficult to apply the model to some objectives. For example, when the model is applied to address model objective #6, it can only calculate at what point in time the mean concentration falls below a target level that no longer poses a human health or ecological risk. This means that roughly (depending on the frequency distribution of the concentrations) half the concentrations are still above the target level. Risk assessment calculations typically depend on the distribution of the concentration and set limits based on a percentage of individuals exceeding a particular concentration. Hence, a frequency distribution of the concentrations is essential. Perhaps it is assumed that the risk assessment calculations can deal with the distribution of the concentration. However, in the application of the model, it is the model that has to generate the distribution of concentrations (since there are no data for the future) and at this point, none of the model components can do this. A comparable argument can be made for the application of the model to model objective #1, i.e. quantify future spatial and temporal distribution of PCBs (both dissolved and particulate forms) within the water column and bed sediment. The mean concentrations that will be produced by the model do not provide information on the statistical distribution of the predicted concentrations. Hence, as the model stands, it is impossible to determine whether any calculated difference in concentration (e.g. as a result of a remediation strategy) is significant and can be treated as a difference in effectiveness among remediation strategies. **(12)**

▶I agree with the EPA that it is not necessary to understand the processes causing small scale variability in concentrations (l.3-4, p.15 Model Validation Report). This is normal is any scientific observation. However, when interpreting the observation or the model calculation (in this case), it is then important to recognize the uncertainty that is associated with the lack of understanding, so that it can be taken into account in the decision analysis. **(2B)**

▶2. The calculation of model uncertainty in the report is seriously flawed. The report includes several references to this. For example, based on findings that the uncertainty in the calculated PCB concentrations is greater than the variability in the sampling data, the authors conclude that "the results should not be interpreted to mean that the uncertainty in model predictions renders the model predictions too uncertain to be usable" (p.5-56 Model Validation Report). The authors further "acknowledge that a true statistical analysis of uncertainty, particularly when uncertainty is propagated through the modeling framework, can produce bounds that may not be possible (or likely) based on an understanding of that system. (p.5-57 Model Validation Report). I think that these are important points. I agree that in models of this complexity, it is virtually impossible to meet some key criteria for a meaningful Monte Carlo Simulation, namely (i) that the model parameters included in the simulations are independent, and (ii) there is insufficient data to properly characterize the variability/uncertainty/error in the many parameters used in the model. Also, it should be recognized that the MCS method does not recognize error in the functional relationships of the model. The KS method is subject to the same limitations as the MCS method and the failure of 4 out of the 56 runs may be due to implausible parameter selections causing the model to crash for these runs due to the problems outlined.

In conclusion, the conditions for a meaningful MCS or KS uncertainty analysis are not met in this study. The calculated uncertainty values are therefore seriously flawed. It is unclear what the meaning of the calculated uncertainty is. The lack of a meaningful uncertainty analysis is a major flaw of the current model because model projections are very difficult, if not impossible to interpret.

To include uncertainty and better characterize it, the modelers could consider using the formidable empirical data set to calculate frequency distributions for the model predicted mean concentrations. To formalize this method, the authors could further develop the MB method described on p. 6-118 (Model Validation Report) by calculating the standard deviation of the mean (i.e. MB). What this will do in the case of PCB concentrations in the sediments is simply project the observed variation in PCB concentrations on the model predicted concentrations. The result is now a distribution of predicted concentration that is grounded in empirical data. This is not a major job, and could be done with little extra work.

Recommendation:

My recommendation is to ignore the MCS and KS results and remove it from the model framework when the model is to be applied in the next phase of the study. Instead I recommend that the modelers use the discrepancy between model predicted concentrations and observed concentrations as a measure of model uncertainty. This is simpler, easier to understand and avoids current computational problems. For example, the data depicted in Figures 6.3.3 to 6.3.8 (Model Validation Report) can provide a reasonable description of the overall model uncertainty. This can be achieved by calculating the confidence limits of the MB used in the report. See Environ. Toxicol. Chem. 23, 2343-2355 for additional details on this method. The application of this method will also generate frequency distributions of model outcomes that can be used in risk assessments. The method that I recommend is not complicated and can be carried out in little time. The main drawback of the application of this method is that it relies on the assumption that the uncertainty identified in past application of the model is a good measure for uncertainty in future model applications. I think that this is a reasonable assumption for some model applications. However, if river functioning is drastically altered by the remediation efforts, this assumption may not apply. **(5B)**

Charge Question #6

Upon review of the model projections of changes in PCB concentrations in environmental media in the example scenarios, are such projection reasonable, using your technical judgment, and are they plausible given the patterns observed in the data.

▶The patterns in the data indicate a slow temporal response of the PCB concentrations in the River. The example scenarios also indicate a small decline of somewhere between 0 to 5% (it is hard to spatially average the concentrations depicted without additional information) over 26 years in response to the assumed loading reductions. I think that this is consistent with the observations.

The data illustrate large small-scale spatial variability in PCB concentrations in sediments. The model appears to capture this to some degree as changes in concentrations relative to the control vary between grid cells within many transects.

I have trouble understanding why PCB concentrations in so many cells of the lower part of the River (Figure 7-1b and 7-3b) increase in concentrations as a result of the loading reductions. This is not what I would expect to happen intuitively given the long history of the PCB contamination problem and the slow response time of PCB concentrations in the river. I would expect concentrations to go down throughout the river, but at higher rates at some locations and lower rates at others.

Comparing scenarios 1 and 2, one would expect that elimination of additional PCB loads in scenario 2 would produce a greater change in PCB concentration over time in scenario 2 than 1. Perhaps, this is the case. It is hard to see from the graphs. However, even if this is not the case, it is possible that concentrations decline over the 26 years are comparable for scenarios 1 and 2. Without more information, it is hard to be more definite.

Based on the current information presented, there is no basis for concluding that the patterns provided in the example scenarios are not plausible given the patterns observed in the data. But more data is needed to support a more positive and definitive conclusion. The example scenarios in the validation report provide little information about the functioning of the model. To address the charge question properly, it is important that the model outcomes of the example scenarios are further analyzed. In particular, it is important to average model outcomes over space and time such that the model predictions can be compared to available data.

Recommendation

I recommend that the behavior of the combined model is explored in greater detail than is presented in the validation report. I recommend that the model outcomes in the model scenarios are aggregated to depict the overall response of the PCB concentrations in the River. This will provide the opportunity to better compare model projections to available data sets and judge whether the model projections are plausible. **(10)**

Charge question #7

Is the final model framework, as calibrated and validated, adequate to achieve the goal of the modeling study to simulate future conditions 1) in the absence of remediation and 2) for use in evaluating the effectiveness of remedial alternatives?

▶The model is the only available tool to simulate the future response of PCB concentrations in response to remediation efforts. The model framework represents a suitable approach to estimating the future time response of PCB concentrations in the River and the calibration and validation of the model have involved significant efforts. The most valuable information for the calibration and validation of the model is a change in PCB concentrations in the River in response to a known reduction in PCB loading. This kind of information was not obtained in the current study as concentrations of PCBs in the River showed little or no significant variation with time. The current model therefore has to rely on the characterization of a number of key state variables for the estimation of the long term temporal response of PCB concentrations to remedial scenarios. The key parameters include the amount of “available” River and floodplains sediments and rates of resuspension, diffusion, bioturbation and subduction. All of these model state variables are either currently unmeasurable or very difficult to measure or estimate. As a result the model’s outcome with regards to the long term time response of PCB concentrations in the River is uncertain. The model uncertainty translates in considerable uncertainty about future PCB concentrations in the River resulting from remediation efforts or the absence of remediation. **(1, 2A)**

▶Any model has inherent uncertainty. So, this is normal. But where the model framework is inadequate is in the recognition of this uncertainty and in the development of adequate methods to characterize or estimate this uncertainty. Without the inclusion of uncertainty in the current model framework, I do not think that it is possible to convincingly distinguish between the effectiveness of different remedial options. **(5B)**

▶The model as it currently stands is incomplete. A lot of excellent work has been done but there are some major gaps that need to be addressed before the model is ready for application. This may sound disappointing to some, especially for those living in the immediate vicinity of the River. However, the remediation options that can be expected to be considered have very large and long lasting impacts on the River and its ecology. Therefore, caution should be exercised and there should be confidence in the outcome of remediation efforts before such remediation takes place. **(12)**

Recommendations:

I have several recommendations for the completion of the model:

1. ▶Include model uncertainty in the model framework and provide guidance about how the results from the uncertainty analysis are to be used when comparing outcomes resulting from different model scenarios.

2. Do not use the MCS or KS method for calculating model uncertainty.

3. Include an uncertainty analysis that takes full advantage of the empirical data that have been collected. As discussed earlier, this can be achieved by comparing observed and model predicted PCB concentrations. **(5B)**

4. ▶Reduce the model run-time from an unacceptable 30 to 50 days to 1 d (at most), such that different model parameterization schemes can be explored for making model projections. **(4)**

5. ▶Conduct sensitivity analyses with the goal to (i) further investigate the parameterization of the sediment-water exchange of PCBs, on which the temporal response of PCB concentrations in the River largely depends and (ii) improve the parameterization of the key processes if possible (see under charge question #1 for additional details on this issue). **(5A)**

6. ▶Continue existing PCB concentration monitoring programs to measure the changes in PCB concentrations over time as a result of the recently completed remediation. Use the data together with calculations of PCB source reductions due to remediation to extend the calibration period and improve the calibration and/or validation of the long term temporal response of the model.

7. Apply a staged and adaptive approach in the planning of River remediation. Plan to gauge the river's response to remedial efforts at certain locations in the River throughout the River remediation. A PCB concentration monitoring program can detect the effect of remedial actions on PCB concentrations over time and space. These data can then be used in the model to further optimize the model, such that the effects of newly planned remedial efforts can be better estimated. **(13)**

Additional Comments

▶DOC-water partitioning

The model assumption that the sorptive capacity of DOC is two orders of magnitude less than that of POC (l.27-28, p.6-67 Model Validation Report) seems out of touch with the literature. The lowest value I have seen to characterize the sorptive capacity of DOC compared to octanol was 0.08, i.e. DOC has 8% of the sorptive capacity of octanol (Burkhard, L.P. Estimating dissolved organic carbon partition coefficients for nonionic organic chemicals. *Environ. Sci. Technol.* 2000, *34* (22), 4663-4668. – Note that in response to p. 2-72 of EPA response, Burkhard states that K_{DOC} = 0.08.Kow, not K_{DOC} = 0.08.K_{POC}). In comparison, the sorptive capacity of POC is approximately 35% of that of octanol (Seth, R.; Mackay, D.; Muncke, J. Estimating the organic carbon partition coefficient and its variability for hydrophobic chemicals. *Environ. Sci. Technol.* 1999, 33 (14), 2390-2394.). Following these papers, the difference in sorptive capacities between DOC and POC is approximately a factor of 35/8 or 4.38. Burkhard et al. did report 95% uncertainty limits of a factor of 20 for the 0.08 value, hence the 0.01 value used in the model is plausible, but it is very low.

Much higher values of the sorptive capacities of DOC in relation to octanol have also been measured. For example, Macintosh et al. report 1.16 (± 0.49) for spiked and 61 (± 47) for native PCBs and phthalates. This compared to 35 (± 19) for POC (Sorption of Phthalate Esters and PCBs in a Marine Ecosystem, Mackintosh et al. Environ. Sci. Technol.; 2006; 40(11); 3481-3488.). The latter results indicate no significant differences between sorptive capacities of DOC and POC and also evidence of DOC-water disequilibria.

The assumed two orders of magnitude difference in sorptive capacities of POC and DOC in the model is, albeit plausible, a very low value. Given the variation in literature data, I recommend that empirical data are used to calibrate this model input requirement. The recent EPA response document suggests that the latter has indeed been done. **(1C)**

▶Diffusive Flux

The importance of a diffusive flux of PCBs from sediments to the water between Holmes Road and New Lennox Road during low flow periods is surprising (p.6-72 Model Validation Report) given the otherwise dominant roles of erosion and deposition (Fig 6.2-62 Model Validation Report). I have never seen a system, especially a riverine system, in which diffusion played such an important role. I think that it was necessary to invoke a high diffusion rate to explain the concentration data. I am not sure if there is any precedent for such a high diffusion rate though. This may be perhaps point to another parameterization problem in the model. I recommend the authors investigate this process in more detail and explore other options for calibrating the model. **(1B)**

Tables

Table 1: Model Objectives

Model objectives:

1. Quantify future spatial and temporal distribution of PCBs (both dissolved and particulate forms) within the water column and bed sediment

2. Quantify historical and current relative contributions of various PCB sources to PCB concentrations in water and bed sediment

3. Quantify historical and current relative contributions of various PCB sources to bioaccumulation in target species

4. Quantify relative risk(s) of extreme storm event(s) contributing to the resuspension of sequestered sediment or the redistribution of PCB-laden sediment in the study area

5. Estimate the time required for PCB-laden sediment to be effectively sequestered by the deposition of uncontaminated material (i.e., natural recovery)

6. Estimate the time required for PCB concentrations in fish tissue to be reduced to levels established during the risk assessment process, that no longer pose either a human health or ecological risk, based upon various response and restoration scenarios

APPENDIX A.6

COMMENTS OF WILBERT LICK

HOUSATONIC RIVER

MODEL VALIDATION

FINAL COMMENTS

EPA FORMAT

Wilbert Lick

University of California

Santa Barbara, California

▶1. In the present model, the processes of sediment erosion, sediment deposition, the finite sorption rate of PCBs by the sediments, and the sediment-water flux due to "diffusion" are described incorrectly and inaccurately. This is exacerbated by a very coarse numerical grid used to define the bathymetry of the river and an unnecessarily fine grid to define the bathymetry/topography of the floodplain. More specific reasons for these comments as well as specific suggestions for improvement of the model are presented in the complete response which is attached. I do not believe that the present model can reasonably account for the relevant processes affecting PCB fate, transport, and bioaccumulation in the Housatonic River to a degree consistent with achieving the goal of the modeling study. **(1, 3A)**

▶2. A good comparison of the model predictions with data is necessary, **but not sufficient**, to evaluate the capability of the model. The low surficial PCB concentrations in Woods Pond and the large variability in PCB concentrations throughout the river are unexplained by the present model. See discussion in the complete response. **(2A, 2B)**

3. Don't know.

▶4. The sensitivity of the model to the parameterization of the significant state and process variables has not been adequately characterized. See discussion in the complete response. **(5A)**

▶5. Because the basic processes are inaccurately described, the uncertainties in model outputs are not correctly acknowledged or described. **(5B)**

▶6. Because the basic processes are inaccurately described and the correct processes may possibly differ by factors of two to ten from those in the present model, the projections of the present model have large potential errors. **(10)**

▶7. The present model is inadequate to achieve the goal of the modeling study to simulate future conditions (1) in the absence of remediation and (2) for use in evaluating the effectiveness of remedial alternatives. **(13)**

July 4, 2006

HOUSATONIC RIVER

MODEL VALIDATION

FINAL COMMENTS

Wilbert Lick
University of California
Santa Barbara, California

▶Introduction

In the long term, the main source of contaminants in the Housatonic River is the bottom sediments (both those in the river and in the floodplain). In this long term, environmental conditions will change with time and will be different than they are at present, especially because of and after remediation. For purposes of predicting water quality, it is therefore essential to accurately determine not only the flux of contaminants between the bottom sediments and the overlying water but also the parameters on which this flux depends. Otherwise, long-term predictions of water quality will not be accurate or believable. Because of this, my comments will emphasize processes which govern the sediment-water flux of contaminants, i.e., sediment erosion, sediment deposition, the sediment-water flux due to “diffusion”, and equilibrium partitioning.

In my last review (Model Calibration, May 13, 2005), I commented extensively on the proposed calibration of both sediment erosion and deposition by means of the measured suspended solids concentration, C. A simple example was given whereby it was easy to see that a numerical model can “predict” the observed values for C with an almost arbitrary value of erosion rate as long as the deposition rate was changed accordingly, i.e., such that the two were equal and gave the observed C. I stated “For a predictive model, the values of erosion rate and deposition rate can not both be determined from calibration of the model by use of the suspended solids concentration alone.” I later stated 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 lead to the incorrect predictions of long-term behavior.”

EPA responded by more-or-less agreeing with the above statements but then stating “This concern does not recognize that there is a constraint imposed by PCB transport that results from resuspension and deposition processes.” Although this argument has some validity, it is not sufficient or correct as I will argue below. This problem of non-unique solutions is important, is more general than that indicated above, and is related to the necessity for accurately determining the basic processes that govern the sediment-water flux of PCBs and other hydrophobic organic chemicals (HOCs). Because of this, I will return to this problem of non-unique solutions after discussing the flux processes mentioned above. Accurate descriptions of the basic processes also depend on an adequate resolution of the bathymetry/topography of the Housatonic. Because of this, comments and suggestions on the problem of numerical gridding in the model will be made. Some discussion on unexplained results of the present model will then be given. A summary and specific suggestions for improvements to the model will conclude my comments. **(1)**

▶Sediment Erosion

In the previous review on Model Calibration (May 13, 2005), I commented extensively on sediment erosion. Although those comments are still valid, I won't repeat all of them here. However, I would like to repeat the following from those comments.

"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."

One 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, sequential 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. I suggested an interpolation procedure that we had used before and which gave us reasonable results. EPA did not seem to have good results with this procedure. Attached is a description of a modified procedure which I have applied to several randomly selected cores on each of the Kalamazoo, Housatonic, and Passaic Rivers. This procedure is more fundamental and correct. In all cases, it produces an n that is equal to two or greater in Eq. (5) above, just as has been demonstrated by all Sedflume laboratory measurements that we have made. With this procedure, the coefficient modifying $(\tau - \tau_{cn})^n$ and n are functions of depth for each representative area core and are obtained from the Sedflume data. The fact that n is two or greater is important in determining erosion rates at high shear stresses, i.e., during big events, and can lead to shear stresses higher by an order of magnitude than the n's chosen by EPA for the Housatonic.

In that previous report, I also stated the following. "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)."

EPA stated that the surficial layer was assumed to be 7 cm thick, not 6 inches as I stated. Since bed coarsening occurs in a layer only a few particle diameters thick, the assumption of a 7 cm mixed layer is also incorrect. **The above comments are still valid. (1A)**

▶Sediment Deposition

I suggested the use of a dynamic flocculation theory that was recently developed and was relatively simple. EPA seems to have had problems implementing the theory. I presume from what they said that this was due to numerical stability problems. That's too bad because it would have given reviewers more confidence in the modeling of deposition rates. **(1A)**

▶The Sediment-Water Flux of HOCs Due to "Diffusion"

The non-erosion/deposition flux of contaminants from the sediments to the overlying water is primarily due to molecular diffusion, bioturbation, and ground-water flow. Each of these processes behaves in a different way and hence needs to be modeled in a different way. EPA has chosen to describe all of these processes by means of a "diffusion" model based on the concept of bioturbation and the assumption of a well-mixed layer. This is the conventional, but not necessarily accurate, approach. It is not accurate simply because the mass transfer approximation (which is not a diffusion approximation) actually used by EPA does not describe or adequately approximate the HOC fluxes of molecular diffusion, bioturbation, or ground-water flow, not even in functional form. The correct functional form (especially its dependence on time) is important because, otherwise, even calibration doesn't work for long term predictions.

EPA did an extensive review of the literature on bioturbation and listed 139 documents of which 43 were retained for detailed review. This listing is somewhat misleading. Of the 43 most relevant documents, almost all are general observations, surveys of the literature, or even surveys of surveys; only about six report quantitative data or laboratory measurements of mixing due to benthic organisms. For example, the figure shown at the last meeting entitled "Bioturbation and Bioavailable Sediment Depths" is from Clarke et al. (2001) and is their interpretation of what organisms do. There is no data (given by Clarke et al. or anywhere else) to support this figure. Clarke et al. is an excellent manuscript, but it is another survey. There is no new data there. None of the documents listed by EPA report on the sediment-water flux or sediment mixing of hydrophobic organic chemicals (HOCs) due to benthic organisms. Since then, EPA has listed additional reports concerned with the flux of HOCs due to benthic organisms. However, these HOCs had relatively low partition coefficients. The only quantitative data that I know of on the effects of benthic organisms on the flux of HOCs with large partition coefficients and their resulting vertical distribution in the sediments is that by Luo et al. (2006) which is attached. Some of my comments are based on this article.

In EPA's modeling, assumptions are (1) a constant (independent of space and time) sediment-water mass transfer coefficient, k, with a value of 1.5 cm/day and (2) a surficial, well-mixed layer whose thickness is constant in time but varies spatially from 4 cm upstream to 7 cm downstream. Measured biomass varies by about a factor of 20 from upstream to downstream.

Benthic mixing is described by EPA in terms of a mixing rate (a diffusion process), also termed a subduction velocity (a convection process), a quantity which I believe is used in the model as a mass transfer coefficient between the sub-surface sediment layers. For the biologically mixed layer, the subduction velocity (values from EPA's table) varies from about 1 x 10^{-9} m/s (1 x 10^{-2} cm/day) upstream to 2 x 10^{-9} m/s (2 x 10^{-2} cm/day) downstream. This factor of two between upstream and downstream seems surprising since the biomass increases by a factor of 20 in the downstream direction. Even more surprising is that the mass transfer coefficient, k, is assumed constant everywhere at 1.5 cm/day. Why doesn't k increase downstream as the biomass increases by a factor of 20?

A surficial well-mixed layer whose thickness is constant in time is assumed in the analysis. An approximate and minimum time for formation of this layer can be calculated from $t = h/v_b$, where h is the thickness of the layer and v_b is the subduction velocity. For the upstream area, h = 4 cm, v_b = 1 x 10^{-9} m/s = 1 x 10^{-2} cm/day, and therefore t = 400 days. For the downstream area, h = 7 cm, v_b = 2 x 10^{-9} m/s = 2 x 10^{-2} cm/day and therefore t = 350 days. In other words, these so-called well-mixed layers are not formed instantaneously and take a minimum of 350 to 400 days to form.

This becomes a little confusing upon examination of the figure presented at the meeting entitled "Contribution to Db from different groups of benthos". Upstream, Db (for all benthos) is approximately 2 x 10^{-3} cm/day while downstream, oligochaetes (the main vertical burrowers and subductionists) contribute a Db of approximately 2.5 x 10^{-2} cm/day. The upstream number is an order of magnitude less than the number in the table cited above as v_b, is probably correct, but gives a time for formation of the well-mixed layer of almost 2000 days (6 years).

Despite the confusion, the numbers for Db are probably correct (to within less than an order of magnitude). This demonstrates (as does Luo et al. more accurately and convincingly) that so-called well-mixed layers for HOCs, if they exist, take a long time (years) to form. Is there any evidence that a well-mixed layer even exists in the Housatonic? I don't believe so.

The mass transfer coefficient, k, for the transport of HOCs by molecular diffusion alone is approximately 1.2 cm/day and decreases slowly with time at a rate which decreases as K_p increases (Deane et al. 1999, Lick et al. 2006, attached). If there is only a small number of organisms, EPA's value of k = 1.5 cm/day compares well with this number. However, with benthic organisms present, Luo et al. give a mass transfer coefficient for HOCs that varies up to 10 cm/day (for benthic organism densities of $10^4/m^2$) and somewhat higher for very dense concentrations of organisms; these values for k are much higher than those that EPA assumes.

No consideration is given to ground-water flow, which can be significant, is a convection and not a diffusion process, and does not involve a well-mixed layer of any sort. **(3D)**

▶Equilibrium Partitioning

After sediment particles are resuspended, they will be transported downstream by the current and eventually settle out of the water column. During this time, the contaminant sorbed to the particle will desorb at some finite rate. The time for a particle to settle out of the water column depends on the settling speed and water depth, while the distance traveled by the particle before depositing depends on the settling time and current speed. For a reasonable range of settling speeds, w, for fine-grained particles/flocs (2×10^{-3} to 1×10^{-1} cm/s) and water depths, h, typical of the Housatonic (1 to 3 m), the settling times ($t = h/w$) are in the range (see table) from 10^3 s (15 min) to 1.5×10^5 s (1.5 days). For medium and coarse grained particles, the settling times are less.

Settling Times (Seconds) for Fine-Grained Particles in the Housatonic

h(cm)	w(cm/s)	
	2×10^{-3}	1×10^{-1}
1×10^2	0.5×10^5 (0.5 days)	1×10^3 (15 min)
3×10^2	1.5×10^5 (1.5 days)	3×10^3 (45 min)

Some experimental results for the adsorption and desorption of HOCs are shown in the appended figures. These sorption times depend on the sediment concentration, particle and floc sizes, conditions of the experiment, and the value of the partition coefficient. The first three figures are for hexachlorobenzene ($K_p = 10^4$ L/kg), while the fourth figure is for the adsorption of HOCs with K_p's from 10^3 to 6.6×10^4 L/kg. The last figure is for the desorption of a PCB with one chlorine (MCB, $K_p = 10^3$ L/kg), hexachlorobenzene (HCB, $K_p = 10^4$ L/kg), and a PCB with six chlorines (HPCB, $K_p = 6.6 \times 10^4$ L/kg, a K_p which is smaller than, but comparable to, the average K_p of about 10^5 L/kg for PCBs in the Housatonic). In 10 days, only about 25% of the HPCB has desorbed; in 50 days, only about 55% has desorbed. As the partition coefficient increases, the amount of desorption in these time intervals will be even less. For PCBs with $K_p = 10^5$ L/kg, the desorption times would be approximately 1.5 times greater than those for HPCB shown here.

Since desorption times are much greater than settling times, it follows that contaminants on resuspended particles will not desorb completely, or even close to completely, in the water column before the particles settle out of the water column. The chemical sorbed to the suspended particles will therefore not reach chemical equilibrium with the chemical dissolved in the water column. It follows that the assumption of equilibrium partitioning is not valid, nor even a good approximation, for the sediments in suspension or in the surficial layers of the bottom sediments.

Incidentally, finite rates of adsorption/desorption (1) will assist in explaining the high observed values of PCBs in the surficial sediments of Woods Pond, (2) the high variability in observed PCB concentrations in the sediments, and (3) probably will force a higher n in erosion formulas (consistent with all other data) since non-equilibrium sorption will not be consistent with EPA's imposed constraint due to equilibrium partitioning.

The conclusion is that finite rates of PCB adsorption and desorption have a major effect on the sediment-water flux due to resuspension/deposition and must therefore be considered in the modeling. **(1C)**

Calibration and Non-Unique Solutions

▶As discussed above, the processes which govern the sediment-water flux of HOCs (sediment erosion, sediment deposition, the sediment-water flux due to "diffusion", and equilibrium partitioning) are described incorrectly and inaccurately. Each of these processes can modify the flux by factors of two to ten. Nevertheless, EPA documents indicate that there is good agreement between the calculated and measured suspended solids concentrations as well as contaminant concentrations. At the same time, sensitivity and uncertainty analyses also seem to say that the model is doing a good job. **How can this be? Are accurate models of sediment and contaminant transport and fate really unnecessary?**

The answers to these questions are in the non-uniqueness of calibrated solutions and the nature of sensitivity and uncertainty analyses. More specifically, it seems that a modeler can assume a wide range of parameters to describe a particular process and still "calibrate" the model so as to determine a mathematical solution which agrees with observations over some time interval. As discussed above, examples of parameters which significantly affect flux processes are (1) different values of the power n in erosion formulas (this gives greatly different erosion rates at high shear stresses depending on the value of n), (2) different parameterizations for settling speeds, (3) different process models and parameters for the sediment-water flux due to "diffusion", and (4) equilibrium partitioning (equivalent to high reaction rates), frozen reaction rates, or anything in between. Calibration of a model does not guarantee that the processes in the model are described properly. At the risk of being repetitive, a water quality modeler can always get good agreement between calculated and observed quantities for a limited time interval and limited conditions, whether the fundamental processes are described properly or not. Another modeler, with quite different descriptions of processes and/or different parameters in his/her model, can get equally good agreement between the calculated and observed quantities. However, future predictions by the different models and modelers will be quite different. This has been demonstrated repeatedly (e.g., see comments on Fox River modeling (Tracy and Keane 2000) in my comments of May 2005). In other words, **calibration is necessary but not sufficient. (1)**

▶As applied to the Housatonic, sensitivity and uncertainty analyses are asking the wrong questions. They do not question whether the basic processes are formulated correctly. As an example, equilibrium partitioning is assumed. Sensitivity and uncertainty analyses never question this assumption, never demonstrate that it is an inaccurate assumption, nor do they propose a suitable reaction rate. **(13)**

▶Since equilibrium partitioning is not valid, a new parameter (the sorption rate) is introduced into the problem. Among other things, this invalidates EPA's statement (when speaking of non-unique solutions) that "there is a constraint imposed by PCB transport that results from resuspension and deposition processes." This constraint, if it exists, is incorrect because non-equilibrium sorption would impose an entirely different constraint than that imposed by equilibrium partitioning. **(1C)**

▶Numerical Gridding

With the present grid, the width of the river is generally approximated as one cell. In the Housatonic, as in most rivers, there are large differences in erosion between the deeper and the shallower parts of the river. Predicting the dissimilar amounts of erosion/deposition across a cross-section of the river is crucial in predicting the long-term exposure of PCBs by erosion and/or natural recovery by deposition. Averaging across the cross-section does not describe the erosion/deposition process accurately. A minimum of three cells across the river (two shallow, near-shore cells and one deeper, center cell) should be used.

In the floodplain, our knowledge of the basic processes of erosion/deposition and the non-erosional/depositional flux is poor. Because of this, a very coarse grid can be used to approximate the processes in this area.

A better description of the bathymetry of the river will increase the computational time. Drastically decreasing the number of grid cells in the floodplain will significantly decrease the computational time. The computational time can also be decreased by (a) separating the hydrodynamics, sediment transport, and contaminant transport calculations and (b) for small and moderate flows, approximate and calculate the hydrodynamics and transport as a sequence of steady-state solutions and only treat big events in detail. **(4)**

Unexplained Results of the Present Model

▶During previous peer review meetings, John List as well as others including myself (but John was most vocal) have emphasized (a) the large unexplained variance in the PCB concentrations in the surficial (six inch) layer of the sediments and (b) the unexplained high concentrations of PCBs in the surficial layers of the sediments in Woods Pond.

EPA had no explanations for these latter two problems, but also stated that an understanding of these problems was not necessary. An understanding of these problems may not be necessary, depending on your point of view, but the problems themselves are quite interesting, deserve some discussion, and are related to the inaccurate modeling of the basic sediment and contaminant flux processes mentioned above. Some discussion of these problems is given here. **(2A, 2B)**

▶In my comments above, I emphasized that PCB sorption times are slow relative to particle settling times and that, because of this, equilibrium partitioning (as assumed in the model) was not a good assumption. Consider the effects of this on PCB transport to Woods Pond. The upstream region of the Housatonic is erosional (on the average), and PCB concentrations near the sediment surface are relatively high (because they were deposited at an earlier time before remediation). As these sediments are eroded, PCBs tend to desorb from the suspended sediments but do not reach anywhere near chemical equilibrium before the sediments are deposited, i.e., the PCB concentrations of depositing sediments are much higher than if they had equilibrated in the overlying water. This erosion/deposition may occur several times before the sediments and their sorbed PCBs reach Woods Pond, or it may only occur once, depending on the flow rate and turbulence. Either way, the sediments deposited in Woods Pond will have higher PCB concentrations than if equilibrium partitioning was assumed, as is observed in field measurements but is not predicted by the present model. **(1C)**

▶As far as the high PCB variance throughout the river is concerned, consider the following. Sediment erosion/deposition depends on the hydrodynamics, e.g., high rates of erosion where the flows are fastest and low rates of erosion (or deposition) where the flows are slow. Because of this, large variations in erosion rates occur across the river (shallow, near-shore areas versus deeper channels in the middle) as well as along the river. This is well illustrated in previous calculations of sediment transport in rivers (Saginaw River (Cardenas et al. 1995), Fox River (Jones et al. 2000)) when a reasonably fine grid was used, i.e., 5 to 11 grid points across the river. Because of the coarse numerical grid, this is not described by the present model. i.e., everything in the model is averaged or smoothed. In some areas (e.g., upstream and/or in the center of the channel where flows are high), mostly erosion occurs and PCB concentrations reflect deposition at an earlier time and are relatively high. These sediments will be transported downstream and will deposit in a non-uniform manner depending on the hydrodynamics. These sediments will retain their high PCB concentrations. In other areas (e.g., sediments from above the upstream boundary, near-shore depositional areas which have received clean sediments from further upstream, or areas where deposition is slow such that surficial sediments can equilibrate with the cleaner overlying water), the PCB concentrations of surficial sediments may be quite low. Slow transport and multiple resuspension/deposition events can also cause low PCB concentrations of surficial sediments.

Because of episodic events, the dependence of erosion/deposition of sediments on highly variable hydrodynamics, the highly variable sources of PCBs (e.g., clean from far upstream, contaminated from deep in the sediments, differences between near-shore, shallow areas and the deeper channel in the center), and of course slow PCB desorption rates (which causes sediments to retain their PCB sorbed concentrations, either high or low), it seems quite plausible that PCB concentrations will be highly variable in space and time in the Housatonic. The present model smooths this all out.

Does this matter? If the only purpose of the model is to duplicate known results, then accurate models of sediment and contaminant transport and fate don't matter. However, if the model is to be used for predictive purposes, then accurate process models do matter. In the predictive mode, future conditions (such as sediment properties, contaminant concentrations in the sediments, concentrations and types of benthic organisms, sediment-water fluxes, flow rates, etc.) will be modified (for example by dredging, capping, or extreme environmental conditions), will change with time, and will be different from those for which the model was calibrated. The basic processes in the present and future are the same. However, their relative effects and significances depend on the modified conditions and will change with time. If the models describing the basic processes have incorrect functional behavior and/or inaccurate parameters, then the model will not predict the long-term behavior properly. Because of this, for the long-term prediction of sediment and contaminant fluxes, it is essential that the functional behavior and parameters of the most significant processes in the model be described correctly. **(3A)**

Summary and Suggestions

▶In the present model, erosion rates are too low (n is too small). During big events, erosion rates may be as large as, or greater than, 10 times that predicted at present. The major effect of this will be on the maximum depth of erosion during big events. A better analysis of Sedflume results as suggested in the attached report will determine a higher and more reasonable value for n. The coefficients in Eqs. (3), (4), and (5) should vary throughout the river as required by the data. **(1A)**

▶In the present model, deposition rates are also too low. Deposition rates are essentially a calibrated parameter and are determined such that the suspended solids concentration is calculated properly. If erosion rates are increased, deposition rates must also be increased in order to maintain good agreement between calculated and observed suspended solids concentrations. **(1A)**

▶Equilibrium partitioning is not a valid assumption. Desorption rates are relatively slow such that, when bottom sediments are resuspended, the sorbed PCBs do not desorb sufficiently rapidly for equilibrium partitioning to be approached before particle deposition occurs. The result is that PCBs sorbed to the suspended solids are not in equilibrium with the PCBs dissolved in the water. The sediment-water flux of PCBs due to resuspension/deposition is therefore much lower than that predicted by equilibrium partitioning, as is assumed in the present model. A finite sorption rate between the PCBs sorbed to the solids and the dissolved PCBs needs to be added to the model and will replace the equilibrium assumption. **(1C)**

▶The sediment-water flux of PCBs due to "diffusion" as calculated by the present model is too low. The formulation of this flux and its parameters are also incorrect. The mass transfer coefficient, k, was given a value of 1.5 cm/day, constant throughout the river;

this value was chosen on the basis of calibration, not on the basis of any field or laboratory measurements.
Since the PCB flux due to resuspension/deposition should be smaller than that predicted by the present model, the PCB flux due to “diffusion” must increase, strictly on a calibration basis, to compensate for this; it must also increase on the basis of a more fundamental investigation of molecular diffusion and bioturbation (Deane et al. 1999, Lick et al. 2006, Luo et al. 2006). In these articles, it is shown that k due to “diffusion” of HOCs has a lower limit of 1.2 cm/day (no organisms) and increases to 10 cm/day (10^4 oligochaetes/m^2) and even greater for larger numbers of organisms, i.e., the sediment-water flux due to “diffusion” may be significantly greater than that predicted by the present model. A realistic value for k should be determined on the basis of the concentrations and types of benthic organisms. When this is done, k will be significantly larger on the average than it is now and will be relatively low upstream but will increase in the downstream direction. **(1B)**

▶The numerical gridding can be improved by increasing the numbr of grid cells across the river and decreasing the number of grid cells in the floosplain. **(3A)** ▶The calculations of the hydrodynamics, sediment transport, and PCB transport should be separated. **(4)**

▶The higher than predicted PCB concentrations in Woods Pond as well as much of the variability seen in the sediment PCB concentrations can be explained by finite PCB sorption rates as well as by the variability in the PCB sources and the hydrodynamics. The improved model should be able to predict some of this and at least suggest the reasons for the remaining variability if finite sorption rates are assumed and a finer grid is adopted. **(2A)**

▶These suggested modifications to the model are relatively simple (except possibly for the re-gridding) and should be able to be accomplished in a year. **(13)**

REFERENCES

Cardenas, M., J. Gailani, C.K. Ziegler, and W. Lick, 1995. Sediment transport in the lower Saginaw River. Marine & Freshwater Res., Vol. 21, pp. 257-274.

Deane, G., Z. Chroneer, and W. Lick, 1999. The diffusion and sorption of hexachlorobenzene in consolidated sediments and saturated soils. J. Envr. Eng., Vol. 125, pp. 689-696.

Jones, C., and W. Lick, 2000. Effects of bed coarsening on sediment transport. Estuarine and Coastal Modeling, VI, pp. 915-930.

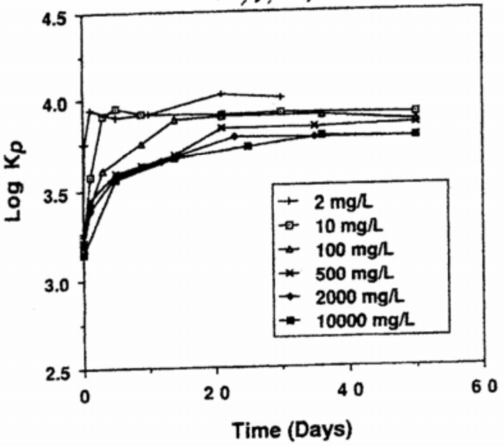
Lick, W., 2006, *Variation, Variability, and Modeling of Sediment Erosion Rates*, Technical Memorandum, University of California, Santa Barbara, CA.

Lick, W., X. Luo, G. Deane, and C. Jones, 2006. "Sediment-water flux of hydrophobic organic chemicals due to molecular diffusion", *Estuarine and Coastal Modeling IX*, ASCE.

Luo, X., W. Lick, and C. Jones, 2006. "Modeling the sediment-water flux of hydrophobic organic chemicals due to bioturbation", *Estuarine and Coastal Modeling IX*, ASCE.

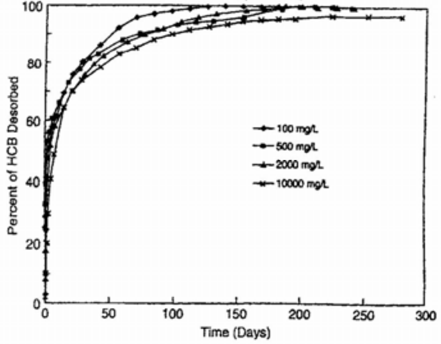
Tracy, J.P., and C.M. Keane, 2000. Peer Review of Models Predicting the Fate and Export of PCBs in the Lower Fox River Below DePere Dam, American Geological Institute, Alexandria, VA.

7.1-1



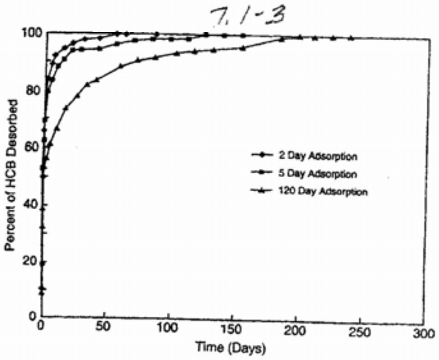
.. Adsorption experiments with hexachlorobenzene (HCB) and pure water. Log K_p as a function of time for sediment concentrations of 2, 10, 100, 500, 2,000, and 10,000 mg/L [8].

7.1-2



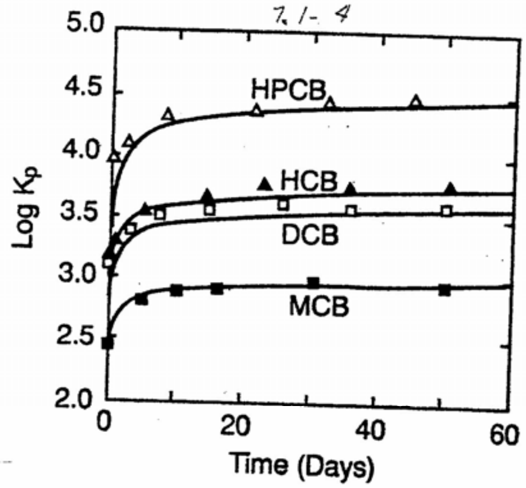
Desorption experiments with hexachlorobenzene (HCB) and pure water. Percent of the initially sorbed HCB which has desorbed is shown as a function of time. Sediment concentrations are 100, 500, 2,000, and 10,000 mg/L [10].

7.1-3



Hexachlorobenzene (HCB) short-term adsorption followed by desorption experiments. The sediment concentration is 500 mg/L. The HCB was first adsorbed to sediments for either 2 d or 5 d and then desorbed. Also shown is long-term adsorption (120 d) followed by desorption, i.e., the standard desorption experiment [10].

7.1-4



Partition coefficients for the adsorption of HCB and three PCB congeners (MCB, DCB, and HPCB) to sediments from the Detroit River. Log K_p is shown as a function of time. Experimental data is shown as open and closed symbols while the modeling results are shown as solid lines.

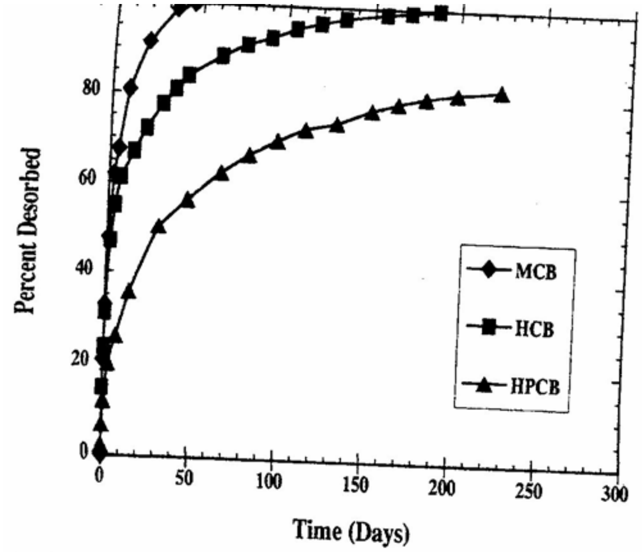


Fig. 8. Desorption experiments with HCB, a monochlorobiphenyl (MCB) and a hexachlorobiphenyl (HPCB). Sediment concentration of 2,000 mg/L.

APPENDIX A.7

COMMENTS OF E. JOHN LIST

FINAL RESPONSE TO QUESTIONS FOR MODEL VALIDATION REPORT

by E. John List, Ph.D., P.E.

Introduction

▶The Peer Review Committee (PRC) has spent much time reviewing the results of the modeling exercises and analyzing the modeling performance. It has devoted a significant level of effort to suggestions as to how to improve the model performance and enhance its predictive capabilities in its application to remedial scenarios. However, to my mind the PRC has also perhaps focused excessively on model omissions and shortcomings and, in the process, has tended to overlook the very real benefits that have come from the modeling exercise, benefits that are substantial, irrespective of the perceived model shortcomings.

First, I believe that the most valuable aspect of the modeling exercise has been the discipline that has been imposed to the assessment of a detailed Housatonic River mass flux balance for the PCBs, total suspended solids (TSS) and water. The key factor in the assessment of any remedial strategy is going to be how it affects these mass fluxes. The modeling exercise enabled a detailed assessment of the relative order of magnitude of the mass fluxes in the various river system components and how these would change under different source flux hypotheses. Although these absolute fluxes may be somewhat in error due to the model shortcomings (to be discussed below), the relative order of magnitude under different hypotheses enables the key leverage points of control for the PCB fluxes to be identified. The available flux data appear to indicate that, although the modeling may have some problems, at least the predicted orders of magnitude appear to be in the range of available data. This suggests that the flux distributions, as described in Figures 1-5 and 1-6, and Figures 1-9 through 1-15, of the document "EPS Responses to Questions from Model Validation Document Overview Meeting", are most probably "in the ball park" and, when properly validated, can be used to develop potential remedial strategies. **(13)**

▶In general terms, the modeling tends to show excellent agreement for water mass fluxes, which is not at all surprising as models to simulate the flow of water in channels have been in use for a long time and a failure on this score would be cause for real alarm. There is only one adjustable parameter for such models and this is the friction factor for the stream, if this parameter is correctly chosen, and the stream cross-sectional area is correct, then the water levels and flow rates can be very accurately simulated. **(3B)** ▶On the other hand, models to simulate the movement of sediment in a fluvial stream are significantly more problematic. The basic difficulty lies in the erodible stream bed. Bed erosion is uniform neither transversally nor longitudinally in the stream. In addition, the bed erosion and the consequent bed form changes result in a change in the stream friction factor, which feeds back into a modification of the hydrodynamics of the flow. **(1A)** ▶If, in addition, the stream bank erodes then long term changes in the direction and shape of the channel can result. **(1D)** ▶Because the PCBs are an integral part of the stream sediments the difficulties of TSS mass balances carry through to the PCBs. To compound the difficulty, the PCB mass flux has two components: one associated with the sediments and the other with the dissolved phase in the water. A further complication is the kinetics of the partitioning process of the PCB between the water and the sediment, which was a matter of some discussion at the Peer Review meeting and may need addressing. **(1C)**

▶EPA and its contractors decided to finesse the problems associated with the non-uniformity of stream bed sediment erosion by using a single computational cell to define the entire width of the river. This means that the bed erosion at any specific river station is defined by the mean bottom shear stress. However, because the actual shear stress on the bed at any location is proportional to the local *depth-averaged* velocity, and the rate of bed erosion (when the shear stress significantly exceeds the critical shear stress) is proportional to the shear stress squared (or an even higher power), it means that the rate of bed erosion in these circumstances is proportional to the fourth power (or even higher) of the mean velocity. Thus, if there is a transverse non-uniform depth-averaged velocity profile *across* the river the rate of bed erosion may vary quite substantially with location on the river cross section. At locations where the depth-averaged velocity is twice that at another location the rate of erosion can be 16 times as great (or even higher). The fact of the matter is that real streams do have widely varying flow velocities across the width of the river and this is the basic reason why the water depth is seldom uniform across the width of the river. It is my professional opinion that it is really inappropriate to attempt to describe the net result of the widely varying rates of erosion that occur in the Housatonic River by using a *cross-sectionally* averaged velocity. The probability that the mean erosion can be parameterized in the model on the basis of cross-sectionally averaged flow properties seems low to me. However, development of a sediment rating curve from the stream data may show that this is possible (see discussion below).

It is acknowledged that it is probably quite impracticable to model the sediment transport and at the same time compute the resulting changes in river bottom topography and their effects on the hydrodynamics. It would seem that a logical approach would be to use a quasi-steady approach in which the stream bed is divided transversally into three or more computational elements that enable at least a partial simulation of the bottom profile and the resulting velocity distribution and bed erosion. An alternative would be to take a longitudinal section of the stream and do a detailed three-dimensional computational fluid dynamics (CFD) analysis of the flow distribution and resulting erosion, then compare the results of this analysis to the one-dimensional approach to determine if a realistic one-dimensional parameterization is at all possible. Perhaps the single best alternative approach to this problem is simply to use empirical sediment and flow rating curves, which is the "old fashioned" empirical way these problems were solved before computers became involved (see, for example, http://www.epa.gov/warsss/sedsource/rivrelat.htm and http://water.usgs.gov/osw/techniques/TSS/Horowitz.pdf, where it is shown that accuracies of ±20% are possible for empirically-derived sediment concentration predictions). Incidentally, the hydraulic modeling could probably also benefit from a comparison of the predicted and actual stage-discharge curves, if this has not already been done. **(3A)**

Apart from this basic problem, which is of particular concern to me, other members of the PRC have other fundamental problems that will only be peripherally referenced here in the specific responses to charge questions that follow.

Question 1:

Considering the changes implemented in the Phase 2 Calibration, does the model as calibrated and validated, based on your technical judgment, reasonably account for the relevant processes affecting PCB fate, transport, and bioaccumulation in the Housatonic River to a degree consistent with achieving the goal of the modeling study?
"i.e., quantify future spatial and temporal distributions of PCBs (both dissolved and particulate forms) within the water column and the bed sediment"

Response:

▶It is my professional opinion that the model does not "reasonably account for the relevant processes", nor, *in its present state*, can it reliably "quantify future spatial and temporal distributions of PCBs".

One of the reasons for this opinion is encompassed in Figure 4.2-79 of the validation report. This figure indicates, from measured data, that 0-6 inch sediment PCB concentrations within the river and flood plain are spread over three orders of magnitude (0.2-200 mg/kg) at locations that are very close together along the river. Furthermore, for almost 12 miles along the river the variance in the distribution of concentrations appears to be fairly uniform. From the fact that the PCB concentrations appear this way after 30-40 years of river action involving sediment erosion and deposition it is reasonable to conclude that the sediment transport processes are not smoothing the sediment PCB concentration distributions. **(2B)**

▶A second reason for this opinion is that recent river bottom survey profiles indicate quite clearly that sediment erosion and deposition is very definitely non-uniform across the river cross-section. By contrast, and as discussed above, the modeling uses a single computational cell to define the entire width of the normal river channel and the local erosion is defined by the fluid shear stress that is averaged across the river cross-section. Because of the non-uniform nature of the actual erosion, and the fact that the model uses an average shear stress to define the erosion, it is highly probable that river bed erosion occurs long before the model predicts its occurrence. Furthermore, the PCB concentration of the sediment that is eroded in the model is the average of the local concentration over that cell. Thus the modeling uses cell averages that reduce the very broad PCB concentration distributions to a local average, and then redeposits that PCB-laden sediment at this average concentration. The net effect of the modeling therefore must be to smooth out the distribution of PCBs, but in reality this has not occurred in the significant length of time in which, according to the model, it should have occurred.

Thus, given the fact that the very broad spectrum of PCB sediment concentrations has been present in the river for more than 30-40 years despite continuous sediment erosion and deposition over these years, it is clear that river erosion and deposition processes actually at work do not smooth the PCB distributions. Absent any explanation of how these PCB concentration distributions are developed and maintained over such a long period of time it seems very unlikely that the modeling will provide any believable answers to remediation strategies. This point has been made by me and others previously. Nevertheless, EPA and its contractors still say they do not know what processes are involved with the PCB fate and transport that would give rise to the very wide concentration spectrums observed. However, it is conceivable that the non-uniform erosion and sediment sorting processes not included in the model are indeed responsible for the non-uniform PCB distributions. **(2B)**

▶A similar point arises in connection with the Validation Report, on Page 4-12 from line 23 et seq. through page 4-13. The high concentrations of PCB in the surface layers of the sediment in Woods Pond (apparent also in Figure 4.2-79) and the lack of an explanation thereof are troubling. Again EPA has thrown up its hands and states that no explanation has been found (see p. 4-13 of the Validation Report). **(2A)**

▶In my professional opinion these two items are really key issues in the fate and transport analysis for the PCB. If neither phenomenon can be explained then how can any model hope to represent what is now occurring or, even more important from the point of view of this project, what will occur in the future. The second item I suspect may be a result of not including wind stresses and wind-induced mixing in Woods Pond. Wind waves continually stir up the sediment and when the wind dies the fine sediment (which apparently is high in PCB) settles last and stays there until the next set of wind waves occurs. The wind waves result in a preferential winnowing of fine material into the surficial sediments of the pond. **(13)**

▶I recognize the fact that these a very difficult problems to address, but if the modeling is to provide believable predictions of the future fate of the PCB in the sediments it seems to me that the processes that have resulted in the existing PCB concentration distributions have to be understood first. **(2A, 2B)**

Question 2:
Are the comparisons of the model predictions with data sufficient to evaluate the capability of the model on the spatial and temporal scales of the final calibration and validation?
Response:
▶The EFDC comparisons between measured and simulated flow and stage offered in Figures 6.2-3 of the Model Validation Report for years 1979-1999 present essentially no measured data and should not be offered up as part of the validation. The comparisons for years 2002-2004 are convincing that EFDC does a good job of representing flow and stage. However, the equivalent time series of Total Suspended Solids (TSS) concentrations for 2002-2004 (Figures 6.2-19 through 6.2-22) do not really provide an adequate basis to conclude that the sediment transport is being properly modeled (which is not surprising considering the points made above). The field data are simply too sparse to make any meaningful comparison between simulation and measured data. Even the data for two specific flow events modeled (Figures 6.2-23 and 6.2-24) offer only a few measured data points for comparison. The use of actual measured sediment rating curves either for calibration or for analysis is urged. The PCB sediment data offered in Figures 6.2-49, 6.2-50, and Figure 6.2-51 indicate just how difficult is the process of validating the PCB fate and transport model and how unlikely it will be that the existing model will provide any real hope of predicting the future fate and transport of the PCB. **(3B)**

Question 3:
Is there evidence of bias in the models, as indicated by the distribution of residuals of model/data comparisons?
Response:
▶With respect to the PCB modeling the results are so tenuous (see Figure 6.2-49) that it is not possible to draw any really solid conclusion with respect to bias. However, an important point is demonstrated by Figure 6.2-50. This figure indicates that the sediment concentrations of PCB in the lower reaches of the project study area have declined somewhat since 1990 (i.e., since remediation activity started). Furthermore, there appears to be a similar decline in the water column concentrations of PCB, as is somewhat evident in the lower panel of Figures 6.2-44. (It is noted that there is no statistical measure of the significance of the decline in either the sediment or the water column, but the data are very suggestive). However, the modeling does not give any indication that any similar rate of decline has occurred (see Figure 6.2-50). In fact the comparison between the data and the model in this figure is strongly suggestive that the model does have a bias. The bias of the model is further reinforced by the results of the hypothetical remedial scenarios plotted in the un-numbered figure on page 56 of the EPA Response to Questions from Model Validation (Document DCN: GE-061406-ADFI), where, despite a reduction in the flux of PCB from the sediments in the upper river reaches, there is no significant change in projected sediment concentration of PCBs in the lower reaches. Thus the model does not match the reduction in sediment concentration that has occurred subsequent to actual post-upstream remediation, and projects no reduction as a result of an hypothetical reduction in the upstream PCB flux, as in the hypothetical examples. This problem needs to be understood and/or fixed before the model is used in real applications. **(2A, 10, 13)**

Question 4:
Have the sensitivities of the models to the parameterization of the significant state and process variables been adequately characterized?
Response:
▶My opinion is that the underlying processes involved with the PCB fate and transport do not appear well enough understood for a meaningful model to be developed. If the processes that seem to define the PCB fate and transport are not included in the model then their sensitivity cannot be assessed. **(1, 5A)**

Question 5:
Are the uncertainties in model output(s) acknowledged and described?
Response:
▶Since the basic processes for the fate and transport of PCB are clearly not properly included in the model their uncertainty cannot be assessed. **(1, 5B)**

Question 6:
Upon review of the model projections of changes in PCB concentrations in environmental media in the example scenarios, are such projections reasonable, using your technical judgment, and are they plausible given the patterns observed in the data?

Response:

▶I am not a biologist, but having said that, to a layperson the broad distributions of the PCB in the biota appear to be reasonably well described, at least at the central tendency level. This may well be due to the fact that partitioning of PCB into biota appears to be closely associated with the spread of concentrations in the environment. **(10, 11)** ▶ In the example scenarios there is no projected effect on the mean PCB concentration in the Woods Pond sediments, which does seem surprising, especially given the results in Figure 6.2-50, where there is a seemingly significant drop in the PCB concentration in the sediment between 1990 and 2000. However, there is a projected rough order of magnitude reduction in the water concentration of PCB in Woods Pond as a result of the hypothetical scenarios. The actual data for Woods Pond referenced above seem to show that there is a reduction in both water column and sediment concentrations coincident with the remediation that has occurred so far. It is not clear why the modeling should necessarily uncouple the water concentration from the sediment concentration in Woods Pond in the hypothetical examples. **(10)**

Question 7:

Is the final model framework, as calibrated and validated, adequate to achieve the goal of the modeling study to simulate future conditions 1) in the absence of remediation and 2) for use in evaluating the effectiveness of remedial alternatives?

Response:

▶I do not think so. For the reasons stated above. **(12)**

Overall Response

".......the Panelist shall describe the alternative approach that, in their opinion, would be sufficient to answer the question and achieve the goal of the modeling study."

▶The primary difficulty here is that the processes that result in the observed concentration distributions of PCBs in the sediment are not known and therefore cannot be included in the model. Second, the geometry of the system is probably too complex to apply properly the sediment transport theory that is known. However, these difficulties should not prevent the development of adequate empirically-based relationships between observed PCB concentrations and fluxes and the water and sediment fluxes. The sediment transport modeling attempts to use a fine scale theory that is applicable to a uniform flow and applies it over an entire river cross-section as if the river had this uniform flow. Many careful observations in the field, of PCB concentrations and fluxes together with stream flow and sediment fluxes, would have provided empirical relationships that would likely have proven much more accurate, useful and predictable. There is nothing wrong with such empirical relationships when the basic theory is unknown or too complex, and in fact EPA and the USGS already use such relationships (see the two websites referenced above). **(13)**