MODEL CALIBRATION: MODELING STUDY OF PCB CONTAMINATION IN THE HOUSATONIC RIVER

PEER REVIEW

Model Calibration
Preliminary Written Comments

Douglas Endicott
Great Lakes Environmental Center
April 22, 2005

Draft Comments on the Housatonic River Model Calibration Report
Douglas Endicott

RESPONSE TO CHARGE FOR THE HYDRODYNAMIC MODELING PEER REVIEW

I. General Overview of Response

I am very impressed by the modeling work that is presented in the calibration report. The project team has done a thorough job in assembling a suite of models that address the major processes affecting PCBs in the Housatonic River, and along the way have overcome a number of obstacles presented by site-specific data that challenge conventional wisdom. I am optimistic that the modeling tools under development here will be valuable in terms of forecasting the outcome of remediation alternatives.

Upon reading the calibration reports ("words were meant to be weighed, not counted") and sitting through the document review meeting, I am convinced that there are a number of issues that must be addressed in order to establish the credibility of the models. Of these, the most important are a number of instances where the model results appear to contradict either the project data or the conceptual model as described in the MFD. First, I have trouble reconciling the 3 to 4 cm maximum change in sediment bed elevation predicted by the sediment transport model for the 14-month calibration period, with the several-foot change based on bathymetric transects and stratigraphic analysis. I realize that spatial aggregation or binning in the model plays a role here. Although aggregation is attractive as an option for dealing with a heterogeneous sediment bed, I am afraid the calibration results fail to demonstrate that the "aggregation approach" works in terms of the way sediment transport processes are described. In other words, how do we know that the highly-nonlinear parameters describing things like cohesive sediment erosion, can be determined by averaging data collected at a limited number of sites?

Another example: over the 14 months, EFDC predicts net loss of PCBs from sediments both upstream (reach 5a) and downstream (reach 6), as well as for the PSA as a whole (20 kg/yr). At the same time, we are told that sediment PCBs are neither being sequestered ("hidden" according to the dictionary) nor have surficial concentrations declined in 25 years[1]. If this is true, then where are the exported PCBs coming from?

The final issue I will throw out under the "credibility" heading is the use of a 1-dimensional model for hydrodynamics and sediment transport. If model grid testing included an evaluation of how bed shear stresses were impacted by the number of lateral segments in the river, the results have not been reported. My semi-informed opinion is that al least 3 lateral segments should be used in a hydrodynamic/sediment transport model of a river. I understand the argument about computational constraints, but I don't think expediency should trump science. When I worked

[1] According to EPAs response to my question E-8, this analysis is presented in 4.4.3.1.3 of the MFD. All I found there was a discussion of sediment data from cores collected in 1996 through 2002. Am I looking in the wrong place? I find it hard to believe there has been no change over 25 years, or that the data exist to prove this!

for the Agency, we were always looking for more optimization and better computer platforms, and investing resources accordingly. Based upon what has been presented to us, it doesn't look like enough effort has gone into exploring these options.

II. Response to Peer Review Questions

In considering the foregoing general issues and evaluating the EPA documents, the Peer Review Panel shall give specific consideration to the following questions. As modeling activities proceed, additional specific questions may be identified the panel to address.

A. ***Modeling Framework and Data Needs***

1. ***Do the modeling frameworks used by EPA include the significant processes affecting PCB fate, transport, and bioaccumulation in the Housatonic River; and are the descriptions of these processes in the modeling framework(s) sufficiently accurate to represent the hydrodynamics, sediment transport, PCB fate and transport, and PCB bioaccumulation in the Housatonic River?***

Neglecting the WWTP, tributaries, and groundwater as PCB sources in the PSA is not justified by the available data. The WWTP may be contributing on the order of a kilogram of PCB (or more) per year. That doesn't sound like a lot in comparison to what's flowing across the confluence currently, but what about after remediation upstream is completed? Same argument applies for tributaries and groundwater.

Neglecting PCB volatilization in EFDC may not be justified. See III (Specific comments on the model calibration report).

2. Based upon the technical judgment of the Peer Review Panel:

a. ***Are the modeling approaches suitable for representing the relevant external force functions (e.g.. hydraulic flows, solids and PCB loads, initial sediment conditions, etc.), describing quantitative relationships among those functions, and developing quantitative relationships between those functions and PCB concentrations in environmental media (e.g., water column, sediments, fish and other biota, etc.)?***

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

b. ***Are the models adequate for describing the interactions between the floodplains and the river?***

No comment.

c. ***Are the models adequate for describing the impacts of rare flood events?***

The calibration report shows that EFDC is capable of predicting the extent of flooding. I am still unsure whether the scour and deposition predictions are reasonable at very high flow rates, and the comparisons to data are problematic due to the spatial resolution of the sediment bed.

d. ***Are the models adequate for discriminating between water-related and sediment-related sources of PCBs to fish and other biota?***

Based upon the calibration report, this is hard to address. No diagnosis of bioaccumulation model results was made to explicitly show the proportion of PCB body burden in different species and reaches contributed by PCB exposure originating in the sediment bed versus the water column. If this discrimination is important, the models should be rerun and this diagnosis made.

3. Again, based upon the technical judgment of the Panel, are the spatial and temporal scales of the modeling approaches adequate to address the principal need for the model - producing sufficiently accurate predictions of the time to attain particular PCB concentrations in environmental media under various scenarios (including natural recovery and different potential active remedial options) to support remedial decision-making in the context described above in the Background section? If not, what levels of spatial and temporal resolutions are required to meet this need?

This question relates to modeling objective #1 from the peer review charge (Quantify future spatial and temporal distribution of PCBs within the water column and bed sediment). I would have to say that EFDC is not representing the observed variability in many sediment properties. Instead, they are being averaged over bins and reaches. The loss of this variability means it cannot be predicted. This doesn't seem to be a problem in terms of forecasting PCB exposure for the bioaccumulation model. It might be a problem if a scenario involved remediating a very localized sediment deposit.

4. Is the level of theoretical rigor of the equations used to describe the various processes affecting PCB fate and transport, such as settling, resuspension, volatilization, biological activity, partitioning, etc., adequate, in your professional judgment, to address the principal need for the model (as defined above)? If not, what processes and what resolution are

required?

Settling and resuspension: I'm sure my esteemed colleagues will address any deficiencies in rigor.

Volatilization: see III (Specific comments on the model calibration report).

Sediment-water exchange: The parameterized mass transfer coefficient (Kf) is very high in comparison to most values I can find in the literature.

Partitioning: Previously, I asked to see the PCB partition coefficients as a function of POC. What I got instead were plots of partition coefficients versus foc (not the same thing). I am still concerned there is an unexplained factor in the water column partitioning data, possible some kind of solids effect.

Metabolism of PCB congeners: Already commented on this.

Coupling of abiotic and biotic PCB exposures: I am not convinced that POC and phytoplankton/ periphyton carbon are interchangeable as sorbents for PCBs. I would like to see the evidence from the PSA that supports this assumption.

PCB elimination by fish: Biphasic elimination is sort of a hidden feature within the QEA foodchain model. How do the elimination rates computed within the model compare to rates measured in fish?

5. What supporting data are required for the calibration/validation of the model on the spatial and temporal scales necessary to address the principal need for the model (as defined above)? What supporting data are required to achieve the necessary level of process resolution in the model?

See response below.

6. Based upon your technical judgment, are the available data, together with the data proposed to be obtained by EPA, adequate for the development of a model that would meet the above referenced purposes? If not, what additional data should be obtained for these purposes?

Ongoing monitoring above confluence: see II.A.2.a

Partitioning data: Since we don't understand what is going on at 15% of the coring locations, maybe it would be appropriate to do some adsorption experiments on the "non-EP" sediments?

Measure things you can measure: PCB concentrations and loads/fluxes in WWTP, tributaries, groundwater … I don't think EPA can justify NOT measuring these potential sources.

Low-flow sediment-water flux: There should be some way to measure bioturbation activity, shouldn't there? Benthic chambers or something…

Sediment mixing: Again, isn't there some experimental approach that could be applied in the field?

B. *Model Calibration*

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

It should be noted that "calibration" in the context of this project has come to mean calibration of short-term (daily to seasonal) changes in model state variables. "Validation" now includes the calibration of long-term (annual to decadal) changes, which are the interesting changes in terms of managing toxic chemicals and making decisions about remedial alternatives. In other words, at this juncture we are unable to evaluate the model's capabilities in terms of its intended application. This limitation has an impact on how thoroughly we can address questions 4, 5 and 6 below.

Net loss of PCBs from Woods Pond contradicts expectation and conceptual model.

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

Bias evident for PCBs across Woods Pond, dissolved PCBs at low flow, PCBs in pore water, and PCB body burdens in fish. Will work on summarizing these concerns (lack of time!).

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

Bank erosion and volatilization omitted from model… should be accounted for.

4. Based upon your technical judgment, have the adequate methodologies been employed to evaluate the sensitivity of the model to descriptions of the relevant processes, and to evaluate the uncertainties of model predictions?

It would be helpful to see the sensitivity of EFDC to an alternative PCB partitioning model (that doesn't ignore non-EP sediments).

At validation stage, surficial sediment bed thickness should be explored in sensitivity analysis.

Analysis of model uncertainty should be expanded to address:

- propagation of uncertainty between models
- uncertainty in statistical models (rating curves) predicting upstream boundary

conditions, as well as uncertainty in flow measurements

5. Is the uncertainty indicated by model-data differences sufficiently inconsequential to permit use of the model to predict differences among remedial options?

This really is a judgment call that depends upon how ambitious the remedial options are. Uncertainty is "inconsequential" if it does not obscure the discrimination between outcomes (i.e., PCB body burdens in fish) of different scenarios. Bioaccumulation model predictions of total PCBs appear to be quite accurate in comparison to estimates based upon National BAFs (www.epa.gov/waterscience/humanhealth/method/tsdvol2.pdf), BSAFs, or predictions made using "generic" food chain models.

6. Are the processes in the model calibrated to the extent necessary for predicting future conditions including future concentrations of PCBs in the environment under natural processes and under potential remedial options for sediments and floodplains soils in the Housatonic River in the reach below the confluence? If not, what additional work needs to be done to calibrate the model?

As I mentioned above (see #1), we can only anticipate the calibration of processes that are influential to long-term model predictions. In this context, I am primarily concerned with the calibration of resuspension fluxes, the sediment-water diffusion flux, and the surficial sediment residence time as defined by the mixed layer thickness. I have already touched on my concerns about the first two processes, so I will focus on the parameterization of the mixed layer thickness.

Willy and EPA agree that the sediment mixed layer depth should be a "non-calibratable parameter". I think this is silly. We need to ask, why is 6 inches (or 3 or 12) the appropriate thickness? What information leads us to that number? I think we need to make use of all lines of evidence when we evaluate this parameter. That should include physical, chemical and biological data as well as what the model can tell us via calibration. Maybe we can agree that EPA has not convincingly demonstrated that the current model parameterization of mixed layer depth is scientifically defensible? If so, then we have some responsibility to suggest ways to improve upon this situation.

If (as EPA suggests) validation does not improve parameterization of sediment-water column interactions, because there has been no trend in surficial sediment PCB concentrations over the past 25 years, then there is a real problem. I guarantee the model will be sensitive to mixed layer depth over decadal-scale forecasts involving a substantial reduction in the upstream PCB boundary condition!

III. Specific Comments on the Model Calibration Report

PCB Fate and Transport Schematic: I usually spend considerable time examining model

schematics, and I find them particularly valuable in understanding how the conceptual model is applied. In this case (Figure B.4-1) there are a number of errors in the schematic which should be corrected. These errors should be obvious to the modeling team.

PCB volatilization: The accuracy of PCB flux estimates being used to justify the neglect of volatilization as a loss process in the PSA, depend upon good values of Henrys constant. The best experimental data I am aware of for this parameter was published by Holly Bamford (Bamford, Poster and Baker, J. Chem Eng. Data, 2000, 45, 1069-1074). She measured Hlc's over a range of temperatures for a bunch of congeners, and also generalized the results into predictions for all PCBs. She found that PCB Henrys constants depended more on the number of ortho-chlorines than on the homolog. If I substitute one of her Hlc values (a representative congener at 18 degrees C) for the value used in the RFI, and repeat the volatilization rate calculation, I get a rate and flux that is about twice as large. I suggest the modeling team evaluate Bamford's data and consider revising the PCB volatilization flux calculation accordingly.

Bioaccumulation of coplanar PCBs: You've seen this one before. The "correction factor" approach used in the bioaccumulation model for coplanar PCBs is not scientifically correct. Based primarily on comparisons of measured PCB BSAFs for fish to model predictions assuming no metabolism (Burkhard et al., ES&T 2004), it appears that congener 77 and possibly congener 126 to a lesser extent are very slightly reduced through metabolism. One cannot rule out a bioavailability effect associated with affinity of more planar PCBs with small amounts of black carbon in water and sediments; however, that is a partitioning issue that was not addressed in EFDC. Since the congener-specific metabolism would probably be aryl hrdrocarbon receptor (AHR) mediated, metabolism in invertebrates is unlikely (i.e., it should only be taking place in the FISH). The first article in the latest ES&T (39/8) is also arguing that metabolism/ biotransformation is the mechanism behind reduced bioaccumulation of specific congeners.