Preliminary Review Regarding Housatonic Model Calibration

E. Adams

April 21, 2005

1. All in all I have been quite impressed and think the models are much further along than they were at our previous review. Following are some general and some specific comments.

2. We are ultimately being asked to evaluate if the models are "good enough". This depends on the problem that they are being asked to address. (Problem identification is usually the first step in model evaluation; see, e.g., Ditmars et al., 1987). Specifically, as I mentioned in Lenox, I believe that model success will vary depending on whether the model is used to assess the relative benefit of different remediation alternatives (e.g., will Scenario A reduce PCB concentrations by 50% compared with a natural attenuation scenario) versus the absolute benefit (e.g., will Scenario A reduce PCB concentrations in fish to acceptable levels by the year 200X). Currently the models are being compared with data in a calibration mode, and it appears that further calibration will take place during the subsequent validation mode. In general a model will perform **better** in a relative assessment, than in a calibration, because model errors tend to cancel. Conversely a model will perform worse in an absolute assessment, than in a calibration, because different data and processes (reflecting remediation options) are involved. A lot was said about model sensitivity. Perhaps the sensitivity should be judged in relation to potential remediation scenarios. That is, identify a couple of scenarios and see how sensitive the models are (in both relative and absolute terms) to variation in uncertain parameters.

3. Of the three models (HSPF, EFDC, and FCM), I sense that EFDC is the most difficult to calibrate because: 1) it is relatively new and has not been used in the current framework (e.g., with both in-channel flows and above bank flows), 2) it is being calibrated over a short period of time (14 months) relative to the time constants of some of the biochemical processes, and 3) compromises are being made because of computational expense. Including data from the earlier years (~1980-1999) will help with the calibration but, as I mentioned, I think it would be nice if the last few years (2000-2004) could be set aside for a true validation (no more parameter tweaking).

4. HSPF has been around for a long time, its developers have had lots of experience with it, and the available data for the Housatonic seems to be on par with (or better than) what is typically available for other sites. There are a lot of semi-empirical parameters that can be adjusted to achieve a good fit and the fits displayed seemed generally acceptable. The biggest errors come when simulating storm events, but this is due to the difficult of getting storms right (both magnitude and timing) with only one hourly rain gauge in the watershed. But this should not be a problem in a statistical sense. (If sediment-laden PCBs are being eroded, we aren't concerned about the exact time of day.) It does seem that the

project might be able to afford a few more hourly rain gauges in the watershed, but I don't think this is a high priority.

5. I don't know much about the FCM, but it seems to be doing pretty well.

6. I am still a bit uncomfortable about the fact that most of the PCB mass is in the floodplain (and hence affected only by relatively rare flood flows), yet the vast majority of the computational time is taken with in-channel flows. Should the same model be used (in the same way) for both? It is not clear whether storms or the routine flows will be most responsible for PCB transport, but the fact that the model can only afford one grid cell over the channel width and a couple of grid cells in the vertical seems to defeat the purpose of a 3D model. This is not to say that the proposed approach is flawed, but merely that the problem is challenging.

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

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

9. The model for the settling velocity of cohesive solids is pretty approximate. The settling process certainly deserves to be included in the sensitivity analysis.

10. There appears to be nominal agreement between the EFDC modeled and the measured fluxes at different cross sections, but these reflect differences between erosion and deposition. Hence error in one process could offset error in the other.

Reference

Ditmars, J.D, E.E. Adams, K.W. Bedford and D.E. Ford (1987). "Performance evaluation of surface water transport and dispersion models" *Journal of Hydraulic Engineering, ASCE*, 113(8): 961-980.